ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, P. ERDŐS, L. KALMÁR, L. RÉDEI,
A. RÉNYI, B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS XI

AKADÉMIAI KIADÓ BUDAPEST

1960

ACTA MATH. HUNG.

INDEX

TOMUS XI

ACTA MATHEMATICA
ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, P. ERDŐS, L. KALMÁR, L. RÉDEI,
A. RÉNYI, B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS XI.

AKADÉMIAI KIADÓ, BUDAPEST
1960

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK
MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST, V., ALKOTMÁNY U. 21

*

Az Acta Mathematica német, angol, francia és orosz nyelven közöl értekezéseket a matematika köréből.

Az Acta Mathematica változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok a következő címre küldendők:

Acta Mathematica, Budapest 502, Postafiók 24.

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Mathematica előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az „Akadémiai Kiadó"-nál (Budapest, V., Alkotmány utca 21. Bankszámla 05-915-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, I., Fő utca 32. Bankszámla 43-790-057-181) vagy annak külföldi képviseleteinél és bizományosainál.

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in deutscher, englischer, französischer und russischer Sprache.

Die Acta Mathematica erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind an folgende Adresse zu senden

Acta Mathematica, Budapest 502, Postafiók 24.

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu richten.

Abonnementspreis pro Band: 110 Forints. Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen „*Kultura*" (Budapest, I., Fő utca 32. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandvertretungen und Kommissionären.

CHARLES JORDAN

1871—1959

CHARLES JORDAN, former professor of the Economic University of Budapest, doctor of the University of Genève, corresponding member of the Hungarian Academy of Sciences, honorary president of the János Bolyai Mathematical Society, member of the International Statistical Institute, honorary ellow of the Royal Statistical Society, fellow of the Institute of Mathematical Statistics and of other scientific societies, died at the age of 88 on 24 December 1959. He contributed a great deal to the development of probability heory, statistics and their applications.

He knew the classical works of the XVII—XIX centuries on probability perhaps better than anybody else. The classical theory of probability received a final touch by his hands; he completed and extended many results of PASCAL, FERMAT, BERNOULLI, MONTMORT, MOIVRE, LAPLACE, POISSON, POINCARÉ tc. His critical study of the classical theory helped the development of modern concepts. His book *"Calculus of finite differences"* published in 1939 contains many important new results (the theory of Stirling's number, of Boole polynomials, extension of the Euler—MacLaurin formula, a new interpolation formula etc.). In a book published in 1956 in Hungarian entitled *"Chapters from he classical theory of probability"* he collected the results of his rich life-work.

Professor MAURICE FRÉCHET has sent an address to the Hungarian Academy of Sciences at the occasion of the death of CHARLES JORDAN from which we reproduce the following words:

"Je viens d'apprendre avec grand regret la mort de CHARLES JORDAN. C'est une grande perte pour la science hongroise et pour la science en général. Je ne connais pas toute l'oeuvre de CHARLES JORDAN... Mais ce que je sais de cette oeuvre suffit pour entraîner mon admiration et provoquer le regret qu'elle n'ait pas attiré toute l'attention qu'elle mérite.... CHARLES JORDAN avait publié un ouvrage sur le Calcul des différences qui était et est encore un véritable trésor de formules et de méthodes utiles.... Enfin, j'ai toujours trouvé auprès de CHARLES JORDAN l'accueil le plus courtois et le plus amical."

CHARLES JORDAN was born on 16 December 1871 in Budapest. He studied at the universities of Paris, Zurich, Manchester and Genève. From 1920 he lectured at the Economic University of Budapest, first as a privat-dozent and later as a professor. A list of his works has been published in the journal *Matematikai Lapok* (7 (1956), pp. 291—294) containing 85 items.

SUR LES APPLICATIONS LINÉAIRES INTÉGRALES DES ESPACES DE FONCTIONS CONTINUES A VALEURS VECTORIELLES

Par

I. SINGER (Bucarest)

(*Présenté par* B. Sz.-Nagy)

1. Soient Q un espace compact, E et F deux espaces de Banach réels, $C(Q, F)$ l'espace de Banach de toutes les fonctions continues $y(\cdot)$ sur Q à aleurs dans F, avec la norme $\|y(\cdot)\| = \max_{q \in Q} \|y(q)\|_F$, et u une application inéaire continue de $C(Q, F)$ dans le dual E' de l'espace E.

D'après [2], définition 7 et proposition 28, l'application u est "*intégrale*", i la forme linéaire U sur[1] $C(Q, F) \otimes E$ définie par

$$(1) \quad U\left[\sum_{i=1}^{n} y_i(\cdot) \otimes x_i\right] = \sum_{i=1}^{n} \langle x_i, u y_i(\cdot) \rangle, \quad y_i(\cdot) \in C(Q, F), \ x_i \in E \quad (i = 1, \ldots, n)$$

st continue pour la norme sur $C(Q, F) \otimes E$ induite par $C(Q, F) \check{\otimes} E$. La "*norme intégrale*" $\|u\|_{\text{int}}$ de l'application u n'est autre que la norme de la orme linéaire U.

D'aprés une définition de [1], l'application u est "*majorée*" s'il existe ur Q une mesure de Radon positive ν, appelée "*majorante*" de u, telle que

$$(2) \qquad \|u y(\cdot)\| \leqq \int_Q \|y(q)\|_F \, d\nu(q) \quad \text{pour toute } y(\cdot) \in C(Q, F).$$

Dans ce cas il existe sur Q une plus petite mesure de Radon positive satisfaisant à (2), désignée par μ et appelée *plus petite majorante de u*.

[1] Rappelons que, si A et B sont deux espaces de Banach, on définit la "norme nductive" $\vee$ sur $A \otimes B$ par

$$|p|_{\vee} = \sup_{\substack{\|a'\| \leqq 1 \\ \|b'\| \leqq 1}} |\langle p, a' \otimes b' \rangle| \qquad (a' \in A', b' \in B'),$$

t qu'on désigne par $A \check{\otimes} B$ le complété de $A \otimes B$ pour la norme $\vee$.

Rappelons aussi, puisque nous l'utiliserons dans ce qui suit, qu'on définit la "norme rojective" $\wedge$ sur $A \otimes B$ par

$$|p|_{\wedge} = \inf \sum_i \|a_i\| \, \|b_i\|,$$

ù l'infimum est pris pour toutes les représentations de p sous la forme $p = \sum_i a_i \otimes b_i$, qu'on désigne par $A \hat{\otimes} B$ le complété de $A \otimes B$ pour la norme $\wedge$.

Pour plus de détails concernant les produits tensoriels normés, voir [2] et [3].

Dans la présente Note nous nous proposons d'examiner les relation entre ces deux classes d'applications linéaires continues ainsi qu'entre l norme intégrale et la norme de la plus petite majorante des application linéaires simultanément majorées et intégrales. Dans le cas particulier, o $F=R=$ la droite numérique munie de la norme habituelle, ce problème été résolu dans la Note antérieure [6], où l'on a démontré qu'une applicatio linéaire intégrale de $C(Q, R)=C(Q)$ dans E' est intégrale si et seulemer si elle est majorée, et qu'alors sa norme intégrale est égale à la norme d sa plus petite majorante. Dans le cas général des applications de $C(Q, F$ dans E', où F est un espace de Banach réel arbitraire, la situation est dif férente. Dans ce cas, le théorème 1 ci-dessous donnera une réponse complèt au problème envisagé.

2. Dans ce qui suit, nous aurons besoin du

LEMME 1. *L'espace $C(Q, F)\check{\otimes}E$ s'identifie à l'espace de Banach $C(Q, F\check{\otimes}E$ de toutes les fonctions continues sur Q, à valeurs dans $F\check{\otimes}E$. Cette identifi cation conserve les normes naturelles.*

DÉMONSTRATION. Désignons par $C(Q)$ l'espace de Banach de toutes le fonctions réelles continues sur Q. En vertu de [2], p. 89—90 et d'une pro priété d'associativité,[2] on a les suivants isomorphismes isométriques:

$$C(Q, F)\check{\otimes}E \equiv (C(Q)\check{\otimes}F)\check{\otimes}E \equiv C(Q)\check{\otimes}(F\check{\otimes}E) \equiv C(Q, F\check{\otimes}E).$$

3. THÉORÈME 1. *Toute application linéaire intégrale u de $C(Q, F)$ dan E' est majorée, et*

(3) $$\|\mu\| \leq \|u\|_{\text{int}}.$$

Pour que la réciproque soit vraie, c'est-à-dire pour que toute application linéair

[2] Cette propriété d'associativité est probablement bien connue; sa démonstration e d'ailleurs facile.

Rappelons qu'à tout $\sum_{i=1}^{n} y_i(\cdot)\otimes x_i \in C(Q, F)\otimes E$ on fait correspondre l'élémer $\left[\sum_{i=1}^{n} y_i(\cdot)\otimes x_i\right](\cdot)$ de $C(Q, F\otimes E)$ défini par

$$\left[\sum_{i=1}^{n} y_i(\cdot)\otimes x_i\right](q) = \sum_{i=1}^{n} y_i(q)\otimes x_i \qquad (q \in Q);$$

pour abréger, nous écrirons $\sum_{i=1}^{n} y_i(\cdot)\otimes x_i$ au lieu de $\left[\sum_{i=1}^{n} y_i(\cdot)\otimes x_i\right](\cdot)$.

Rappelons aussi, puisque nous l'utiliserons plus tard, qu'une identification analogu est aussi valable pour $L^1(Q, \nu, F)\otimes E$ et $L^1(Q, \nu, F\otimes E)$.

majorée de $C(Q, F)$ dans E' soit intégrale, il faut et il suffit que l'application linéaire canonique J de $F\hat{\otimes}E$ dans $F\check{\otimes}E$ soit un isomorphisme vectoriel topologique du premier espace sur le second, et dans ce cas on a

$$\|u\|_{\text{int}} \leqq \|J^{-1}\| \, \|\mu\|. \tag{4}$$

DÉMONSTRATION. 1° Soit u une application linéaire intégrale de $C(Q, F)$ dans E'. Alors la forme linéaire U définie par (1) est continue sur $C(Q, F)\otimes E$ pour la norme induite par $C(Q, F)\check{\otimes}E$, donc $\bar{U}$, son prolongement par continuité à $C(Q, F)\check{\otimes}E$, c'est-à-dire (lemme 1) à $C(Q, F\check{\otimes}E)$, est majorée.[3] Il existe donc sur Q une mesure de Radon positive ν telle que

$$|\bar{U}[y(\cdot)\otimes x]| \leqq \int_Q \|y(q)\| \cdot \|x\| d\nu(q), \quad y(\cdot) \in C(Q, F), \quad x \in E$$

(puisque $\|y(q)\otimes x\|_{F\check{\otimes}E} = \|y(q)\|_F \|x\|_E$ pour chaque $q \in Q$), donc (formule (1)) telle que

$$|\langle x, uy(\cdot)\rangle| \leqq \|x\| \int_Q \|y(q)\| d\nu(q), \quad y(\cdot) \in C(Q, F), \quad x \in E,$$

d'où

$$\|uy(\cdot)\| = \sup_{\|x\|\leqq 1} |\langle x, uy(\cdot)\rangle| \leqq \int_Q \|y(q)\| d\nu(q), \quad y(\cdot) \in C(Q, F),$$

c'est-à-dire l'application u est majorée par ν.

Il s'ensuit aussi que toute majorante de $\bar{U}$ est une majorante de u. Par conséquent, si μ est la plus petite majorante de u et ω la plus petite majorante de $\bar{U}$, on a $\mu \leqq \omega$, d'où

$$\|\mu\| \leqq \|\omega\| = \|\bar{U}\| = \|U\| = \|u\|_{\text{int}},$$

ce qui n'est autre que l'inégalité (3).

2° Soit u une application linéaire majorée de $C(Q, F)$ dans E' et soit ν une majorante de u. Alors, en vertu de (2), on peut prolonger par continuité u en une application linéaire continue $\tilde{u}$ de l'espace $L^1(Q, \nu, F)$ des

[3] Rappelons (voir p. ex. [6]) qu'une forme linéaire Φ sur $C(Q, F\check{\otimes}E)$ est continue si et seulement si elle est majorée; on dit que Φ est majorée s'il existe sur Q une mesure de Radon positive ν telle que

$$|\Phi[X(\cdot)]| \leqq \int_Q \|X(q)\|_{F\check{\otimes}E} d\nu(q) \quad \text{pour toute} \quad X(\cdot) \in C(Q, F\check{\otimes}E).$$

De plus, dans ce cas on a $\|\Phi\| = \|\omega\|$, où ω désigne la plus petite majorante de Φ.

fonctions à valeurs dans F et absolument sommables pour ν, muni de la norme $\|z(\cdot)\|_1 = \int_Q \|z(q)\|_F d\nu(q)$, dans E', et l'on aura

$$(2') \qquad \|\tilde{u}z(\cdot)\| \leqq \|z(\cdot)\|_1 \quad \text{pour toute} \quad z(\cdot) \in L^1(Q, \nu, F).$$

Prolongeons la forme linéaire U sur $C(Q, F) \otimes E$ définie par (1), en une forme linéaire $\tilde{U}$ sur $L^1(Q, \nu, F) \otimes E$, en posant par définition

$$(5) \qquad \tilde{U}\left[\sum_{i=1}^{n} z_i(\cdot) \otimes x_i\right] = \sum_{i=1}^{n} \langle x_i, \tilde{u} z_i(\cdot)\rangle, \quad z_i(\cdot) \in L^1(Q, \nu, F), \quad x_i \in E \quad (i = 1, \ldots, n).$$

Comme $L^1(Q,\nu,F) \hat{\otimes} E \equiv [L^1(Q,\nu) \hat{\otimes} F] \hat{\otimes} E \equiv L^1(Q,\nu) \hat{\otimes} (F \hat{\otimes} E) \equiv L^1(Q,\nu,F \hat{\otimes} E)$ (voir [2], p. 59, théorème 2 et p. 51, fin du §1), par (5) et (2') on aura alors, pour toute "fonction simple" $Z(\cdot) = \sum_{i=1}^{n} \varphi_{e_i}(\cdot) w_i$, où

$$w_i = \sum_{k=1}^{m_i} y_k^{(i)} \otimes x_k^{(i)} \in F \otimes E \quad (y_k^{(i)} \in F, x_k^{(i)} \in E;\ k = 1, \ldots, m;\ i = 1, \ldots, n)$$

et les $e_i \subset Q$ sont des ensembles boréliens deux-à-deux sans point commun tels que $\bigcup_{i=1}^{n} e_i = Q$, et où φ_{e_i} désigne la fonction caractéristique de e_i,

$$|\tilde{U}[Z(\cdot)]| = \left|\tilde{U}\left[\sum_{i=1}^{n} \sum_{k=1}^{m_i} \varphi_{e_i}(\cdot) y_k^{(i)} \otimes x_k^{(i)}\right]\right| =$$

$$= \left|\sum_{i=1}^{n} \sum_{k=1}^{m_i} \langle x_k^{(i)}, \tilde{u}[\varphi_{e_i}(\cdot) y_k^{(i)}]\rangle\right| \leqq \sum_{i=1}^{n} \sum_{k=1}^{m_i} |\langle x_k^{(i)}, \tilde{u}[\varphi_{e_i}(\cdot) y_k^{(i)}]\rangle| \leqq$$

$$\leqq \sum_{i=1}^{n} \sum_{k=1}^{m_i} \|x_k^{(i)}\| \, \|\tilde{u}[\varphi_{e_i}(\cdot) y_k^{(i)}]\| \leqq \sum_{i=1}^{n} \sum_{k=1}^{m_i} \|x_k^{(i)}\| \int_Q \|\varphi_{e_i}(q) y_k^{(i)}\| d\nu(q) =$$

$$= \int_Q \sum_{i=1}^{n} \varphi_{e_i}(q) \sum_{k=1}^{m_i} \|y_k^{(i)}\| \, \|x_k^{(i)}\| d\nu(q).$$

Or, comme $\|w_i\|_{F \hat{\otimes} E} = \inf \sum_{k=1}^{m_i} \|y_k^{(i)}\| \, \|x_k^{(i)}\|$, le inf étant pris pour toutes les expressions $\sum_{k=1}^{m_i} y_k^{(i)} \otimes x_k^{(i)} \in F \otimes E$ telles que $w_i = \sum_{k=1}^{m_i} y_k^{(i)} \otimes x_k^{(i)}$, et comme les

$e_i \subset Q$ sont deux-à-deux sans point commun, il s'ensuit que

$$|\tilde{U}[Z(\cdot)]| \leqq \int_Q \sum_{i=1}^n \varphi_{e_i}(q) \|w_i\|_{F \hat{\otimes} E}\, d\nu(q) = \int_Q \|Z(q)\|_{F \hat{\otimes} E}\, d\nu(q).^4$$

Par conséquent, la forme linéaire $\tilde{U}$ définie par (5) est bornée sur un sous-espace vectoriel partout dense[5] de $L^1(Q, \nu, F \hat{\hat{\otimes}} E)$ pour la norme induite par

[4] En effet, si l'on avait $|\tilde{U}[Z(\cdot)]| > \int_Q \sum_{i=1}^n \varphi_{e_i}(q) \|w_i\|_{F \hat{\otimes} E}\, d\nu(q)$, il existerait un nombre $\varepsilon > 0$ tel que

$$(*) \qquad |\tilde{U}[Z(\cdot)]| > \varepsilon + \int_Q \sum_{i=1}^n \varphi_{e_i}(q) \|w_i\|_{F \hat{\otimes} E}\, d\nu(q).$$

D'autre part, il existe des expressions $\sum_{k=1}^{m_i} y_k^{(i)} \otimes x_k^{(i)} = w_i$ $(i=1,\dots,n)$ telles que $\|w_i\|_{F \hat{\otimes} E} + \frac{\varepsilon}{\|\nu\|} > \sum_{k=1}^{m_i} \|y_k^{(i)}\| \|x_k^{(i)}\|$. $(i=1,\dots,n)$, d'où

$$|\tilde{U}[Z(\cdot)]| \leqq \int_Q \sum_{i=1}^n \varphi_{e_i}(q) \sum_{k=1}^{m_i} \|y_k^{(i)}\| \|x_k^{(i)}\|\, d\nu(q) \leqq$$

$$\leqq \int_Q \sum_{i=1}^n \varphi_{e_i}(q) \left(\|w_i\|_{F \hat{\otimes} E} + \frac{\varepsilon}{\|\nu\|} \right) d\nu(q) = \int_Q \sum_{i=1}^n \varphi_{e_i}(q) \|w_i\|_{F \hat{\otimes} E}\, d\nu(q) + \varepsilon,$$

contrairement à (*).

[5] Soit, en effet, $Z(\cdot) \in L^1(Q, \nu, F)$ quelconque et soit $\varepsilon > 0$. On sait bien qu'on peut alors trouver une $\sum_{i=1}^n \varphi_{e_i}(\cdot) h_i$, où les $e_i \subset Q$ sont des ensembles boréliens deux-à-deux sans point commun tels que $\bigcup_{i=1}^n e_i = Q$ et où $h_i \in F \hat{\otimes} E$ $(i=1,\dots,n)$ telle que

$$\int_Q \|Z(q) - \sum_{i=1}^n \varphi_{e_i}(q) h_i\|_{F \hat{\otimes} E}\, d\nu(q) < \frac{\varepsilon}{2}.$$

D'autre part, on peut trouver des $w_i \in F \otimes E$ tels que

$$\|h_i - w_i\|_{F \hat{\otimes} E} < \frac{\varepsilon}{2\|\nu\|} \qquad (i = 1, \dots, n).$$

Il s'ensuit, compte tenu de $\sum_{i=1}^n \varphi_{e_i}(q) = 1$, l'inégalité désirée

$$\int_Q \|Z(q) - \sum_{i=1}^n \varphi_{e_i}(q) w_i\|_{F \hat{\otimes} E}\, d\nu(q) < \varepsilon.$$

$L^1(Q, \nu, F\hat{\otimes}E)$. On peut donc prolonger $\tilde{U}$ par continuité à $L^1(Q, \nu, F\hat{\otimes}E)$ entier et on aura, en désignant ce prolongement encore par $\tilde{U}$,

$$|\tilde{U}[Z(\cdot)]| \leqq \int_Q \|Z(q)\|_{F\hat{\otimes}E} d\nu(q) \quad \text{pour toute} \quad Z(\cdot) \in L^1(Q, \nu, F\hat{\otimes}E),$$

donc, en particulier, pour toute $Z(\cdot) \in C(Q, F\hat{\otimes}E)$.

Supposons maintenant que l'application linéaire canonique J de $F\hat{\otimes}E$ dans $F\check{\otimes}E$ soit un isomorphisme vectoriel topologique du premier espace sur le second. Alors l'inégalité ci-dessus entraîne

$$|\tilde{U}[X(\cdot)]| \leqq \|J^{-1}\| \int_Q \|X(q)\|_{F\check{\otimes}E} d\nu(q) \quad \text{pour toute} \quad X(\cdot) \in C(Q, F\check{\otimes}E).$$

Or, cela prouve, en considérant la mesure $\nu' = \|J^{-1}\|\nu$, que la forme linéaire U sur $C(Q, F)\otimes E$ définie par (1) peut être prolongée en une forme linéaire *majorée* $\tilde{U}$ sur $C(Q, F\check{\otimes}E)$, donc en une forme linéaire *continue* sur $C(Q, F\check{\otimes}E)$, c'est-à-dire (lemme 1), en une forme linéaire continue sur $C(Q, F)\check{\otimes}E$. Par conséquent, l'application linéaire u de $C(Q, F)$ dans E' est intégrale.

Il s'ensuit aussi qu'on a, pour toute majorante ν de u et toute $y_i(\cdot) \in C(Q, F)$, $x_i \in E$ $(i = 1, \ldots, n)$,

$$\left| U\left[\sum_{i=1}^{n} y_i(\cdot)\otimes x_i\right] \right| \leqq \|J^{-1}\| \int_Q \left\| \sum_{i=1}^{n} y_i(q)\otimes x_i \right\|_{F\check{\otimes}E} d\nu(q) \leqq$$

$$\leqq \|J^{-1}\| \, \|\nu\| \max_{q\in Q} \left\| \sum_{i=1}^{n} y_i(q)\otimes x_i \right\|_{F\check{\otimes}E},$$

d'où, compte tenu du lemme 1,

$$\|u\|_{\text{int}} = \|U\| \leqq \|J^{-1}\| \, \|\mu\|,$$

ce qui n'est autre que l'inégalité (4).

3° Supposons enfin que toute application linéaire majorée de $C(Q, F)$ dans E' soit intégrale. Nous montrerons qu'alors les ensembles $(F\hat{\otimes}E)'$ et $(F\check{\otimes}E)'$ sont identiques. En effet, comme en vertu de l'interprétation de ces duals on a $(F\check{\otimes}E)' \subset (F\hat{\otimes}E)$, on n'a qu'à montrer l'inclusion opposée $(F\hat{\otimes}E)' \subset (F\check{\otimes}E)'$. Supposons le contraire, c'est-à-dire qu'il existe une $\bar{f} \in (F\hat{\otimes}E)'$, $\bar{f} \notin (F\check{\otimes}E)'$. Cela signifie que l'application linéaire de F dans E' qui correspond à $\bar{f}$ et que nous désignerons par f, est continue mais non intégrale. Nous disons qu'alors l'application linéaire continue u de $C(Q; F)$

dans E', définie par

$$uy(\cdot) = f[y(q_0)], \quad y(\cdot) \in C(Q, F), \tag{6}$$

où q_0 est un point fixé (d'ailleurs quelconque) de Q, est majorée, mais non intégrale, contrairement à notre hypothèse. En effet, d'une part

$$\|uy(\cdot)\|_{E'} = \|f[y(q_0)]\|_{E'} \leqq \|f\| \, \|y(q_0)\|_F = \int_Q \|y(q)\|_F \, d\nu(q),$$

où $\nu = \|f\| \varepsilon_{q_0} =$ la masse $+\|f\|$ dans le point q_0; donc u est majorée. D'autre part, comme f est une application linéaire non intégrale de F dans E', on peut trouver (voir [2], définition 7 et proposition 28) une suite d'éléments $\sum_{i=1}^{n_k} y_i^{(k)} \otimes x_i^{(k)}$ $(k=1, 2, \ldots)$ de $F \otimes E$ telle que

$$\left| \sum_{i=1}^{n_k} \langle x_i^{(k)}, f[y_i^{(k)}(q_0)] \rangle \right| \geqq k \left\| \sum_{i=1}^{n_k} y_i^{(k)} \otimes x_i^{(k)} \right\|_{F \check{\otimes} E} \quad (k=1, 2, \ldots). \tag{7}$$

Or, pour les fonctions $y_i^{(k)}(\cdot) \in C(Q, F)$ définies par

$$y_i^{(k)}(q) \equiv y_i, \quad q \in Q \quad (i=1, \ldots, n_k; k=1, 2, \ldots),$$

on aura alors, en vertu de (6) et (7),

$$\left| U\left[\sum_{i=1}^{n_k} y_i^{(k)}(\cdot) \otimes x_i^{(k)} \right] \right| = \left| \sum_{i=1}^{n_k} \langle x_i^{(k)}, u y_i^{(k)}(\cdot) \rangle \right| =$$
$$= \left| \sum_{i=1}^{n_k} \langle x_i^{(k)}, f[y_i^{(k)}(q_0)] \rangle \right| \geqq k \max_{q \in Q} \left\| \sum_{i=1}^{n_k} y_i^{(k)}(q) \otimes x_i^{(k)} \right\|_{F \check{\otimes} E}$$
$$(k=1, 2, \ldots),$$

donc, compte tenu du lemme 1, la forme linéaire U définie par (1) est non bornée sur $C(Q, F) \otimes E$ pour la norme induite par $C(Q, F) \check{\otimes} E$; c'est-à-dire, l'application linéaire u est non intégrale. Cela achève la démonstration de l'égalité $(F \hat{\otimes} E)' = (F \check{\otimes} E)'$. Or, en vertu de [2], lemme 15, cette égalité entraîne que l'application linéaire canonique J de $F \hat{\otimes} E$ dans $F \check{\otimes} E$ est un isomorphisme vectoriel topologique du premier espace sur le second.

Le théorème 1 est démontré.

REMARQUE 1. Dans le cas particulier où $F = R =$ la droite numérique munie de la norme habituelle, l'application linéaire canonique J de $F \hat{\otimes} E$ dans $F \check{\otimes} E$ est un isomorphisme isométrique (donc $\|J\| = 1$) du premier espace sur le second, et on retrouve du théorème 1 ci-dessus le théorème 1 de la Note [6].

Remarque 2. Dans la partie 3° de la démonstration ci-dessus nous avons donné, en passant, une méthode de construire, chaque fois où la condition de la deuxième partie du théorème 1 n'est pas satisfaite, un exemple simple d'une application linéaire majorée non intégrale u de $C(Q, F)$ dans E'.

Remarque 3. Si F ou E est de dimension finie, la condition de la deuxième partie du théorème 1, c'est-à-dire la condition que l'application linéaire canonique J de $F \hat{\otimes} E$ dans $F \check{\otimes} E$ soit un isomorphisme sur, est toujours satisfaite. Réciproquement, d'après une hypothèse de Grothendieck ([2], p. 153, ou [4], § 6), il est possible que cette condition ne soit satisfaite que si F ou E est de dimension finie; en tout cas, c'est vrai pour certains espaces de Banach E ou F concrets (voir [2], chap. I, § 4, no. 5). D'ailleurs, en vertu du théorème 1, l'hypothèse ci-dessus de Grothendieck serait démontrée, si l'on pouvait construire, pour chaque couple E, F d'espaces de Banach de dimension infinie, une application linéaire majorée non intégrale u de $C(Q, F)$ dans E'.

4. La démonstration ci-dessus du théorème 1 s'appuie directement sur la définition des applications linéaires intégrales, resp. majorées. En utilisant les théorèmes de représentation intégrale de telles applications, on peut donner une autre démonstration du théorème 1, qui semble être plus courte (toutefois, la méthode précédente ne perd pas son intérêt; voir la remarque faite, à ce propos, dans [6]).

Notamment, on a le suivant théorème de représentation intégrale:[6]

(A) *Pour qu'une application linéaire continue u de $C(Q, F)$ dans E' soit intégrale, il faut et il suffit qu'il existe une fonction unique f_e définie sur les ensembles boréliens $e \subset Q$, à valeurs dans l'espace de toutes les applications linéaires intégrales de F dans E' muni de la norme intégrale, complètement additive, à variation bornée et régulière, telle que*

$$uy(\cdot) = \int_Q \langle y(q), df_q \rangle, \quad y(\cdot) \in C(Q, F) \tag{8}$$

(intégrale de Radon—Gowurin). *Dans ce cas, on a*

$$\|u\|_{\text{int}} = \sup \sum_{i=1}^{n} \|f_{e_i}\|_{\text{int}}, \tag{9}$$

où le sup *est pris pour toutes les partitions finies de Q en ensembles boréliens disjoints e_i.*

[6] Remarquons que dans le cas où $F \check{\otimes} E$ est séparable, on peut aussi obtenir un autre théorème de représentation intégrale des applications linéaires intégrales de $C(Q, F)$ dans E', à l'aide d'un théorème de représentation de Grothendieck ([2], proposition 32) et du lemme 1.

Démonstration. Soit u une application linéaire intégrale de $C(Q,F)$ dans E'. Alors la forme linéaire U définie par (1) est continue sur $C(Q,F)\otimes E$ pour la norme induite par $C(Q,F)\check{\otimes}E$, donc, en prolongeant U par continuité à $C(Q,F)\check{\otimes}E$ et compte tenu du lemme 1, $U\in[C(Q,F\check{\otimes}E)]'$. En vertu du [5], théorème 1, il existe donc une fonction unique $\bar{f}_e$ définie sur les ensembles boréliens $e\subset Q$, à valeurs dans $(F\check{\otimes}E)'$, complètement additive, à variation bornée et régulière, telle que l'on ait, pour tout $x\in E$, $\varphi_i(\cdot)\in C(Q)$, $y_i\in F$ $(i=1,\dots,n)$,

$$\left\langle x, u\left[\sum_{i=1}^{n}\varphi_i(\cdot)y_i\right]\right\rangle = U\left[\sum_{i=1}^{n}\varphi_i(\cdot)y_i\otimes x\right] = \sum_{i=1}^{n}\int_Q \varphi_i(q)\,d\langle y_i\otimes x, \bar{f}_q\rangle,$$

et que

$$\|u\|_{\text{int}} = \|U\| = \underset{e\subset Q}{\text{Var}}\,\bar{f}_e = \sup\sum_{i=1}^{n}\|\bar{f}_{e_i}\|, \tag{10}$$

où le sup est pris pour toutes les partitions finies de Q en ensembles boréliens disjoints e_i.

Donc, en désignant pour chaque ensemble borélien $e\subset Q$ par f_e l'application linéaire intégrale de F dans E' qui correspond à $\bar{f}_e$, c'est-à-dire l'application f_e définie par

$$\langle x, f_e(y)\rangle = \langle y\otimes x, \bar{f}_e\rangle,\quad x\in E,\quad y\in F,$$

la première des quantités ci-dessus sera égale à[7]

$$\sum_{i=1}^{n}\int_Q \varphi_i(q)\,d\langle x, f_q(y_i)\rangle = \left[\sum_{i=1}^{n}\int_Q \varphi_i(q)\,df_q(y_i)\right](x) = \left(\int_Q\left\langle\sum_{i=1}^{n}\varphi_i(q)y_i, df_q\right\rangle\right)(x).$$

Il s'ensuit que

$$u\left[\sum_{i=1}^{n}\varphi_i(\cdot)y_i\right] = \int_Q\left\langle\sum_{i=1}^{n}\varphi_i(\cdot)y_i, df_q\right\rangle,\quad \varphi_i(\cdot)\in C(Q),\quad y_i\in F\quad(i=1,\dots,n),$$

d'où, compte tenu du fait que l'ensemble des $\sum_{i=1}^{n}\varphi_i(\cdot)y_i$ est dense dans $C(Q,F)$, de la continuité de u et du fait que f_e est à variation bornée,

$$u[y(\cdot)] = \int_Q\langle y(q), df_q\rangle\quad\text{pour toute}\quad y(\cdot)\in C(Q,F).$$

[7] Les deux dernières égalités sont immédiates en vertu des définitions des trois intégrales qui y figurent.

En vertu de (10) et de l'égalité $\|\bar{f_e}\| = \|f_e\|_{\text{int}}$, on aura aussi (9). Enfin, f_e est évidemment complètement additive, régulière, et univoquement déterminée par u.

Soit, réciproquement, f_e une fonction définie sur les ensembles boréliens $e \subset Q$, à valeurs dans l'espace de toutes les applications linéaires intégrales de F dans E' muni de la norme intégrale, complètement additive, à variation bornée et régulière. L'application u de $C(Q, F)$ dans E' définie par (8) est alors linéaire sur l'espace $C(Q, F)$; l'inégalité

$$\left| \int_Q \langle y(q), df_q \rangle \right| \leqq \max_{q \in Q} \|y(q)\| \sup \sum_{i=1}^{n} \|f_{e_i}\|_{\text{int}}, \quad y(\cdot) \in C(Q, F),$$

(où le sup est pris pour toutes les partitions finies de Q en ensembles boréliens disjoints e_i) entraîne qu'elle est aussi bornée. En vertu de ce qui précède, f_e est la seule fonction d'ensemble complètement additive, à variation bornée et régulière satisfaisant à (8) et l'on a

$$\|u\|_{\text{int}} = \sup \sum_{i=1}^{n} \|f_{e_i}\|_{\text{int}}$$

(où le sup est pris comme ci-dessus), ce qui achève la démonstration.

D'autre part, on a aussi ([1], théorème 2):

(B) *Pour qu'une application linéaire continue u de $C(Q, F)$ dans E' soit majorée, il faut et il suffit qu'il existe une fonction unique f_e définie sur les ensembles boréliens $e \subset Q$, à valeurs dans l'espace de toutes les applications linéaires continues de F dans E' muni de la norme habituelle, complètement additive, à variation bornée et régulière, telle que*

$$uy(\cdot) = \int_Q \langle y(q), df_q \rangle, \quad y(\cdot) \in C(Q, F)$$

(intégrale de Radon—Gowurin). *Dans ce cas on a, pour la plus petite majorante μ de u,*

$$\|\mu\| = \operatorname*{Var}_{e \subset Q} f_e = \sup \sum_{i=1}^{n} \|f_{e_i}\|,$$

où le sup est pris pour toutes les partitions finies de Q en ensembles boréliens disjoints e_i.

Cela étant, on peut démontrer théorème 1 comme suit:

1°° Comme toute application linéaire intégrale f de F dans E' est continue et comme $\|f\| \leqq \|f\|_{\text{int}}$, (A) et (B) entraînent que toute application linéaire intégrale u de $C(Q, F)$ dans E' est majorée et que $\|\mu\| \leqq \|u\|_{\text{int}}$.

2°° Si l'application linéaire canonique J de $F\hat{\otimes}E$ dans $F\check{\otimes}E$ est un isomorphisme vectoriel topologique du premier espace sur le second, sa transposée ${}^tJ$ est un isomorphisme de $(F\check{\otimes}E)'$ *sur* $F(\hat{\otimes}E)'$, donc toute application linéaire continue f de F dans E' est alors intégrale et $\|f\|_{\text{int}} \leqq \|({}^tJ)^{-1}\| \, \|f\| = \|J^{-1}\| \, \|f\|$. Par conséquent, en vertu de (B) et (A) toute application linéaire majorée de $C(Q, F)$ dans E' est alors intégrale, et $\|u\|_{\text{int}} \leqq \|J^{-1}\| \cdot \|\mu\|$.

3°° Si toute application linéaire majorée de $C(Q, F)$ dans E' est intégrale, on a $(F\hat{\otimes}E)' = (F\check{\otimes}E)'$. En effet, s'il existait une $\bar{f} \in (F\hat{\otimes}E)'$, $\bar{f} \notin (F\check{\otimes}E)'$, l'application linéaire continue u de $C(Q, F)$ dans E' définie par (6) serait, en vertu de (B) et (A) et compte tenu de $f[y(q_0)] = \int_Q \langle y(q), df_e \rangle$, où

$$f_e = \begin{cases} f & \text{pour} \quad e \ni q_0, \\ 0 & \text{pour} \quad e \not\ni q_0, \end{cases}$$

majorée mais non intégrale, contrairement à l'hypothèse. Or, en vertu de [2], lemme 15, l'égalité $(F\hat{\otimes}E)' = (F\check{\otimes}E)'$ entraîne que l'application linéaire canonique J de $F\hat{\otimes}E$ dans $F\check{\otimes}E$ est un isomorphisme vectoriel topologique du premier espace sur le second, ce qui achève la démonstration.

(Reçu le 27 mai 1959.)

Bibliographie

[1] N. Dinculeanu, Mesures vectorielles et opérations linéaires, *Comptes Rendus Acad. Sci. Paris*, **246** (1958), p. 2328—2331.

[2] A. Grothendieck, Produits tensoriels topologiques et espaces nucléaires, *Memoirs of the Amer Math. Soc.*, **16** (1955), p. 1—191, 1—140.

[3] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques, *Boletim da Soc. Mat. São Paulo*, **8** (1956), p. 1—79.

[4] A. Grothendieck, Sur certaines classes de suites dans les espaces de Banach, et le théorème de Dvoretzky—Rogers, *Boletim da Soc. Mat. São Paulo*, **8** (1956), p. 83—110.

[5] И. Зингер, Линейные функционалы на пространстве непрерывных отображений бикомпактного хаусдорфого пространства в пространство Банаха, *Revue de math. pures et appl.*, **2** (1957), p. 301—315.

[6] I. Singer, Sur les applications linéaires intégrales des espaces de fonctions continues. I, *Revue de math. pures et appl.*, **4** (1959), p. 391—401.

SOME FINITE SUMMATION FORMULAS OF ARITHMETIC CHARACTER. II

By

L. CARLITZ (Durham, USA)

(*Presented by* P. TURÁN)

1. Generalizing some results of MIKOLÁS [2] and MORDELL [3], the writer [1] has proved the following theorem:

Let $f_1(x), \dots, f_n(x)$ be functions of x of period 1 that satisfy the equation

$$\sum_{r=0}^{k-1} f_i\left(x+\frac{r}{k}\right) = C_i^{(k)} f_i(kx) \qquad (i=1, \dots, n), \tag{1.1}$$

for arbitrary integral $k \geqq 1$, and where $C_i^{(k)}$ is independent of x. Also let $a_1, \dots, a_n$ be positive integers that are relatively prime in pairs and put $A = a_1 a_2 \dots a_n$. Then we have

$$\sum_{r=0}^{kA-1} f_1\left(x_1+\frac{r}{a_1 k}\right) \dots f_n\left(x_n+\frac{r}{a_n k}\right) = \tag{1.2}$$

$$= C_1^{(a_1)} \dots C_n^{(a_n)} \sum_{r=0}^{k-1} f_1\left(a_1 x_1 + \frac{r}{k}\right) \dots f_n\left(a_n x_n + \frac{r}{k}\right),$$

where $x_1, \dots, x_n$ are arbitrary.

The functional equation (1.1) is suggested by the multiplication formula for the Bernoulli polynomials and functions. The Euler polynomials $E_m(x)$ and the corresponding functions $\bar{E}_m(x)$ defined by

$$\bar{E}_m(x) = E_m(x) \quad (0 \leqq x < 1), \quad \bar{E}_m(x+1) = -\bar{E}_m(x),$$

on the other hand, satisfy

$$\sum_{r=0}^{k-1} (-1)^r E_m\left(x+\frac{r}{k}\right) = k^{-m} E_m(kx), \tag{1.3}$$

where now k is restricted to odd values. More generally, if we put

$$\frac{1-\zeta}{1-\zeta e^t} e^{xt} = \sum_{m=0}^{\infty} \Phi_m(x, \zeta) \frac{t^m}{m!} \qquad (\zeta \neq 1)$$

and assume that ζ is a primitive e-th root of unity, $e > 1$, then $\Phi_m(x, \zeta)$

satisfies

$$(1.4) \qquad \sum_{r=0}^{k-1} \zeta^r \Phi_m\left(x+\frac{r}{k}, \zeta\right) = k^{-m} \Phi_m(kx, \zeta),$$

provided $k \equiv 1 \pmod e$. For $e=2$, $\Phi_m(x, -1) = E_m(x)$ and (1.4) reduces to (1.3). If we define $\overline{\Phi}_m(x, \zeta)$ by means of

$$(1.5) \qquad \overline{\Phi}_m(x, \zeta) = \Phi_m(x, \zeta) \quad (0 \leqq x < 1), \quad \overline{\Phi}_m(x+1, \zeta) = \zeta^{-1} \overline{\Phi}_m(x, \zeta),$$

then $\overline{\Phi}_m(x, \zeta)$ also satisfies (1.4).

In view of the above it seems natural to consider functions $f(x, \zeta)$ that satisfy

$$(1.6) \qquad f(x+1, \zeta) = \zeta^{-1} f(x, \zeta)$$

and

$$(1.7) \qquad \sum_{r=0}^{k-1} \zeta^r f\left(x+\frac{r}{k}, \zeta\right) = C^{(k)} f(kx, \zeta),$$

where ζ is a primitive e-th root of unity, $C^{(k)}$ is independent of x and $k \equiv 1 \pmod e$. For $e=1$, $\zeta=1$, (1.7) reduces to (1.1); for $e=2$, $\zeta=-1$, $f(x, -1) = \overline{E}_m(x)$, (1.7) reduces to (1.3), while for $e \geqq 2$, $f(x, \zeta) = \Phi_m(x, \zeta)$, (1.7) reduces to (1.4). We shall prove the following theorem:

THEOREM 1. *Let $n \geqq 1$; $e_1, \ldots, e_n \geqq 1$; let ζ_i denote a primitive e_i-th root of unity. Let $f_i(x, \zeta_i)$ be functions that satisfy*

$$(1.8) \qquad f_i(x+1, \zeta_i) = \zeta_i^{-1} f_i(x, \zeta_i) \qquad (i=1, \ldots, n)$$

and

$$(1.9) \qquad \sum_{r=0}^{k-1} \zeta_i^r f_i\left(x+\frac{r}{k}, \zeta_i\right) = C_i^{(k)} f_i(kx, \zeta_i),$$

where $C_i^{(k)}$ is independent of x and

$$(1.10) \qquad k \equiv 1 \pmod e,$$

where e is the least common multiple of $e_1, \ldots, e_n$. Also let $a_1, \ldots, a_n$ be positive integers that are relatively prime in pairs and such that

$$(1.11) \qquad a_i \equiv 1 \pmod e \qquad (i=1, \ldots, n);$$

put $A = a_1 a_2 \ldots a_n$. Then we have

$$(1.12) \qquad \sum_{r=0}^{kA-1} \zeta_1^r \zeta_2^r \ldots \zeta_n^r f_1\left(x_1+\frac{r}{a_1 k}, \zeta_1\right) \ldots f_n\left(x_n+\frac{r}{a_n k}, \zeta_n\right) =$$
$$= C \sum_{r=0}^{k-1} \zeta_1^r \zeta_2^r \ldots \zeta_n^r f_1\left(a_1 x_1+\frac{r}{k}, \zeta_1\right) \ldots f_n\left(a_n x_n+\frac{r}{k}, \zeta_n\right),$$

where

(1. 13) $$C = C_1^{(a_1)} C_2^{(a_2)} \dots C_n^{(a_n)}.$$

2. We remark that in the left member of (1. 12) the sum may be extended over any complete residue system (mod kA), while in the right member the sum may be extended over any complete residue system (mod k). Indeed, if we replace r by $r+kA$ on the left, we get, using (1. 8), (1. 10), (1. 11),

$$(\zeta_1 \dots \zeta_n)^{r+kA} f_1\left(x_1 + \frac{r+kA}{a_1 k}, \zeta_1\right) \dots f_n\left(x_n + \frac{r+kA}{a_n k}, \zeta_n\right) =$$

$$= (\zeta_1 \dots \zeta_n)^{r+1} f_1\left(x_1 + \frac{r}{a_1 k} + \frac{A}{a_1}, \zeta_1\right) \dots f_n\left(x_n + \frac{r}{a_n k} + \frac{A}{a_n}, \zeta_n\right) =$$

$$= (\zeta_1 \dots \zeta_n)^{r+1} \zeta_1^{-A/a_1} \dots \zeta_n^{-A/a_n} f_1\left(x_1 + \frac{r}{a_1 k}, \zeta_1\right) \dots f_n\left(x_n + \frac{r}{a_n k}, \zeta_n\right) =$$

$$= (\zeta_1 \dots \zeta_n)^r f_1\left(x_1 + \frac{r}{a_1 k}, \zeta_1\right) \dots f_n\left(x_n + \frac{r}{a_n k}, \zeta_n\right).$$

On the right side, if we replace r by $r+k$, we get

$$(\zeta_1 \dots \zeta_n)^{r+k} f_1\left(x_1 + \frac{r+k}{k}, \zeta_1\right) \dots f_n\left(x_n + \frac{r+k}{k}, \zeta_n\right) =$$

$$= (\zeta_1 \dots \zeta_n)^r f_1\left(x_1 + \frac{r}{k}, \zeta_1\right) \dots f_n\left(x_n + \frac{r}{k}, \zeta_n\right).$$

To prove the theorem it will be convenient to modify somewhat the method of proof used in [1]. To begin with we replace the summation index n the left member of (1. 12) by $ks+t$, where s runs through a complete residue system (mod A) and t a complete residue system (mod k). Thus, if S stands for the left side of (1. 12), it is evident from (1. 10) that

(2. 1) $$S = \sum_{t=0}^{k-1} \zeta_1^t \dots \zeta_n^t \sum_{s \pmod A} \zeta_1^s \dots \zeta_n^s f_1\left(x_1 + \frac{s}{a_1} + \frac{t}{a_1 k}, \zeta_1\right) \dots$$
$$\dots f_n\left(x_n + \frac{s}{a_n} + \frac{t}{a_n k}, \zeta_n\right).$$

Put

$$A = a_1 a_2 \dots a_n = a_i A_i \qquad (i = 1, \dots, n);$$

since the a_i are relatively prime in pairs, it follows that $(A_1, A_2, \dots, A_n) = 1$. Thus we may put

(2. 2) $$s = A_1 s_1 + A_2 s_2 + \dots + A_n s_n;$$

when s runs through a complete residue system (mod A), each s_i will run through a complete residue system (mod a_i) and conversely. We have by (2. 2)

$$\frac{A_1 s_1 + \cdots + A_n s_n}{a_i} = \frac{s - A_i s_i}{a_i} + \frac{A_i s_i}{a_i};$$

since $a_i | A_j$ for $i \neq j$, it is clear that the first fraction on the right is an integer which by (1. 11) is congruent to $s - s_i \pmod{e_i}$. Applying (1. 8) we get

$$f_i\left(x_i + \frac{s}{a_i} + \frac{t}{a_i k}, \zeta_i\right) = \zeta_i^{-s+s_i} f_i\left(x_i + \frac{a_i s_i}{a_i} + \frac{t}{a_i k}, \zeta_i\right).$$

Consequently, (2. 1) becomes

$$S = \sum_{t=0}^{k-1} \zeta_1^t \ldots \zeta_n^t \sum_{s_i \pmod{a_i}} \zeta_1^{s_1} \ldots \zeta_n^{s_n} f_1\left(x_1 + \frac{A_1 s_1}{a_1} + \frac{t}{a_1 k}, \zeta_1\right) \ldots$$

$$\ldots f_n\left(x_n + \frac{A_n s_n}{a_n} + \frac{t}{a_n k}, \zeta_n\right) = \sum_{t=0}^{k-1} \zeta_1^t \ldots \zeta_n^t \prod_{i=1}^{n} \sum_{s_i \pmod{a_i}} \zeta_i^{s_i} f_i\left(x_i + \frac{A_i s_i}{a_i} + \frac{t}{a_i k}, \zeta_i\right).$$

Now the inner sum is evidently equal to

$$\sum_{s_i \pmod{a_i}} \zeta_i^{A_i s_i} f_i\left(x_i + \frac{A_i s_i}{a_i} + \frac{t}{a_i k}, \zeta_i\right) =$$

$$= \sum_{r_i \pmod{a_i}} \zeta_i^{r_i} f_i\left(x_i + \frac{r_i}{a_i} + \frac{t}{a_i k}, \zeta_i\right) = C_i^{(a_i)} f_i\left(a_i x_i + \frac{t}{k}, \zeta_i\right),$$

if we observe that in the left member of (1. 9) the summation may be extended over any complete residue system (mod k). Hence we get

$$S = C_1^{(a_1)} \ldots C_n^{(a_n)} \sum_{t=0}^{k-1} \zeta_1^t \ldots \zeta_n^t f_1\left(a_1 x_1 + \frac{t}{k}, \zeta_1\right) \ldots f_n\left(a_n x_n + \frac{t}{k}, \zeta_n\right).$$

This evidently completes the proof of (1. 12).

3. Some special cases of Theorem 1 may be noted. Since (1. 10) is the only condition on k, we may in particular take $k = 1$. Then (1. 12) reduces to

(3. 1) $$\sum_{r=0}^{A-1} \zeta_1^r \zeta_2^r \ldots \zeta_n^r f_1\left(x_1 + \frac{r}{a_1}, \zeta_1\right) \ldots f_n\left(x_n + \frac{r}{a_n}, \zeta_n\right) =$$

$$= C f_1(a_1 x_1, \zeta_1) \ldots f_n(a_n x_n, \zeta_n),$$

subject to the condition stated in Theorem 1. This result may be thought of as an immediate generalization of (1. 7).

For $f_i(x, \zeta_i) = \Phi_{m_i}(x, \zeta_i)$, (1. 12) becomes

(3. 2) $$\sum_{r=0}^{Ak-1} \zeta_1^r \zeta_2^r \dots \zeta_n^r \bar{\Phi}_{m_1}\left(x_1 + \frac{r}{a_1 k}, \zeta_1\right) \dots \bar{\Phi}_{m_n}\left(x_n + \frac{r}{a_n k}, \zeta_n\right) =$$
$$= C \sum_{r=0}^{k-1} \zeta_1^r \zeta_2^r \dots \zeta_n^r \bar{\Phi}_{m_1}\left(a_1 x_1 + \frac{r}{k}, \zeta_1\right) \dots \bar{\Phi}_{m_n}\left(a_n x_n + \frac{r}{k}, \zeta_n\right),$$

where

$$C = a_1^{-m_1} a_2^{-m_2} \dots a_n^{-m_n}.$$

4. The formula

(4. 1) $$\sum_{r=0}^{k-1} (-1)^r B_m\left(x + \frac{r}{k}\right) = -\frac{m}{2k^{m-1}} E_{m-1}(kx) \qquad (k \text{ even})$$

relating the Euler to the Bernoulli polynomials, as well as the more general relation

(4. 2) $$\sum_{r=0}^{k-1} \zeta^r \bar{B}_m\left(x + \frac{r}{k}\right) = \frac{m}{(\zeta - 1) k^{m-1}} \bar{\Phi}_{m-1}(kx, \zeta),$$

where $\zeta^k = 1$, $\zeta \neq 1$, suggest the possibility of a further extension of (1. 12). Let $f(x, \zeta)$ be a function that satisfies (1. 6) and (1. 7), where ζ is a primitive e-th root of unity. Let c be a fixed integer $\equiv 1 \pmod{e}$ and let η be a ce-th root of unity such that $\zeta = \eta^c$. Now put

(4. 3) $$\sum_{r=0}^{c-1} \eta^r f\left(x + \frac{r}{c}, \zeta\right) = K^{(c)} g(cx, \eta)$$

where $K^{(c)}$ is independent of x. Then in the first place

$$K^{(c)} \eta g(cx + 1, \eta) = \sum_{r=0}^{c-1} \eta^{r+1} f\left(x + \frac{r+1}{c}, \zeta\right) =$$
$$= \sum_{r=0}^{c-1} \eta^r f\left(x + \frac{r}{c}\, \zeta\right) + \eta^c f(x+1, \zeta) - f(x, \zeta).$$

Using (1. 6) this reduces to

(4. 4) $$g(x+1, \eta) = \eta^{-1} g(x, \eta).$$

In the next place, if $k \equiv 1 \pmod{ce}$,

$$K^{(c)} \sum_{r=0}^{k-1} \eta^r g\left(x + \frac{r}{k}, \eta\right) = \sum_{r=0}^{k-1} \eta^r \sum_{s=0}^{c-1} \eta^s f\left(\frac{x}{c} + \frac{r}{ck} + \frac{s}{c}, \zeta\right) =$$
$$= \sum_{r=0}^{k-1} \sum_{s=0}^{c-1} \eta^{r+sk} f\left(\frac{x}{c} + \frac{r+sk}{ck}, \zeta\right) = \sum_{t=0}^{ck-1} \eta^t f\left(\frac{x}{c} + \frac{t}{ck}, \zeta\right) =$$
$$= \sum_{r=0}^{k-1} \sum_{s=0}^{c-1} \eta^{cr+s} f\left(\frac{x}{c} + \frac{r}{k} + \frac{s}{ck}, \zeta\right) = \sum_{s=0}^{c-1} \eta^s \sum_{r=0}^{k-1} \zeta^r f\left(\frac{x}{c} + \frac{r}{k} + \frac{s}{ck}, \zeta\right) =$$
$$= C^{(k)} \sum_{s=0}^{c-1} \eta^s f\left(\frac{kx}{c} + \frac{s}{c}, \zeta\right) = K^{(c)} C^{(k)} g(kx, \eta).$$

Assuming $K^{(c)} \neq 0$, we have therefore proved

$$\sum_{r=0}^{k-1} \eta^r g\left(x+\frac{r}{k}, \eta\right) = C^{(k)} g(kx, \eta) \qquad (k \equiv 1 \pmod{ce}), \tag{4.5}$$

where $C^{(k)}$ is defined by (1.7).

We have incidentally proved that

$$\sum_{r=0}^{ck-1} \eta^r f\left(x+\frac{r}{ck}, \zeta\right) = K^{(c)} C^{(k)} g(ckx, \eta) \qquad (k \equiv 1 \pmod{e}). \tag{4.6}$$

It is accordingly convenient to put

$$K^{(ck)} = K^{(c)} C^{(k)}. \tag{4.7}$$

We remark that, by (1.6) and (4.4), the summations occurring in (4.3), (4.5), (4.6) may, in each instance, be extended over arbitrary complete residue systems.

In (4.3) η may be replaced by $\omega\eta$, where ω is any c-th root of unity. It then follows easily that

$$f(x, \zeta) = \frac{K^{(c)}}{c} \sum_{\omega} g(cx, \omega\eta), \tag{4.8}$$

the summation extending over all c-th roots of unity. Indeed, from (4.6) we get the somewhat more general result

$$\sum_{r=0}^{k-1} f\left(x+\frac{r}{k}, \zeta\right) = \frac{K^{(ck)}}{c} \sum_{\omega} g(ckx, \omega\eta) \qquad (k \equiv 1 \pmod{e}), \tag{4.9}$$

which reduces to (4.7) for $k=1$.

5. We shall now prove the following theorem:

THEOREM 2. *Let $k \geqq 1$; $n \geqq 1$; $e_1, \ldots, e_n \geqq 1$; $c_1, \ldots, c_n \geqq 1$. Let ζ_i be a primitive e_i-th rooth of unity and let η_i be a $c_i e_i$-th root of unity such that $\zeta_i = \eta_i^{c_i}$. Let $f_i(x, \zeta_i)$ be functions that satisfy* (1.8) *and* (1.9) *and put*

$$\sum_{r=0}^{c_i-1} \eta_i^r f_i\left(x+\frac{r}{c_i}, \zeta_i\right) = K^{(c_i)} g_i(c_i x, \eta_i) \qquad (i = 1, \ldots, n). \tag{5.1}$$

Let $a_1, \ldots, a_n$ be positive integers that are relatively prime in pairs and such that

$$a_i | c_i, \quad a_i \equiv 1 \pmod{ce} \qquad (i = 1, \ldots, n), \tag{5.2}$$

where c is the least common multiple of $c_1, \ldots, c_n$ and e is the least common multiple of $e_1, \ldots, e_n$. Then if $A = a_1 a_2 \ldots a_n$ and

$$k \equiv 1 \pmod{ce}, \tag{5.3}$$

we have

$$\sum_{r=0}^{Ak-1} \eta_1^r \dots \eta_n^r f_1\left(x_1+\frac{r}{a_1 k}, \zeta_1\right) \dots f_n\left(x_n+\frac{r}{a_n k}, \zeta_n\right)= \tag{5.4}$$
$$=K \sum_{r=0}^{k-1} \eta_1^r \dots \eta_n^r g_1\left(a_1 x_1+\frac{r}{k}, \eta_1\right) \dots g_n\left(a_n x_n+\frac{r}{k}, \eta_n\right),$$

where

$$K=K^{(a_1)} K^{(a_2)} \dots K^{(a_n)}. \tag{5.5}$$

We remark first that the summations in (5. 4) may be extended over arbitrary residue systems and secondly that it follows from (4. 6) and (5. 1) that

$$\sum_{r=0}^{a_i-1} \eta_i^r f_i\left(x+\frac{r}{a_i}, \zeta_i\right)=K^{(a_i)} g_i(a_i x, \eta_i) \qquad (i=1, \dots, n). \tag{5.6}$$

The proof of the theorem is similar to the proof of Theorem 1. We first replace r by $ks+t$, so that, if S stands for the left member of (5. 4), we have by (5. 3)

$$S=\sum_{t=0}^{k-1} \eta_1^t \dots \eta_n^t \sum_{s \,(\mathrm{mod}\, A)} \eta_1^s \dots \eta_n^s f_1\left(x_1+\frac{s}{a_1}+\frac{t}{a_1 k}, \zeta_1\right) \dots f_n\left(x_n+\frac{s}{a_n}+\frac{t}{a_n k}, \zeta_n\right).$$

In the inner sum we employ (2. 2); we remark that by (5. 2)

$$s \equiv A s_i \equiv s_i \pmod{ce} \qquad (i=1, \dots, n).$$

Consequently, the inner sum becomes

$$\sum_{s_i \,(\mathrm{mod}\, a_i)} \prod_{i=1}^{n} \eta_i^{s_i} f_i\left(x_i+\frac{s}{a_i}+\frac{t}{a_i k}, \zeta_i\right)=\prod_{i=1}^{n} K^{(a_i)} g_i\left(a_i x+\frac{t}{k}, \eta_i\right),$$

by (5. 6). The theorem now follows at once.

6. If we take $c_1=\dots=c_n$, then $\eta_i=\zeta_i$ and Theorem 2 reduces to Theorem 1. In the second place for $k=1$, (5. 4) becomes

$$\sum_{r=0}^{A-1} \eta_1^r \dots \eta_n^r f_1\left(x_1+\frac{r}{a_1}, \zeta_1\right) \dots f_n\left(x_n+\frac{r}{a_n}, \zeta_n\right)= \tag{6.1}$$
$$=K g_1(a_1 x_1, \eta_1) \dots g_n(a_n x_n, \eta_n),$$

subject to the hypothesis of the theorem.

If we put

$$f(x, \zeta)=\overline{\Phi}_m(x, \zeta), \quad g(x, \eta)=\overline{\Phi}_m(x, \eta)$$

and recall that

$$(6.2)\qquad \sum_{r=0}^{a-1} \eta^r \bar{\Phi}_m\left(x+\frac{r}{c}, \zeta\right) = \frac{1-\zeta}{1-\eta} c^{-m} \bar{\Phi}_m(cx, \eta),$$

where $\zeta = \eta^c$, then (4. 3) is satisfied with

$$K^{(c)} = \frac{1-\zeta}{1-\eta} c^{-m}.$$

We according obtain the following special case of (5. 4):

$$(6.3)\qquad \sum_{r=0}^{Ak-1} \eta_1^r \dots \eta_n^r \bar{\Phi}_{m_1}\left(x_1+\frac{r}{a_1 k}, \zeta_1\right) \dots \bar{\Phi}_{m_n}\left(x_n+\frac{r}{a_n k}, \zeta_n\right) =$$

$$= K \sum_{r=0}^{k-1} \eta_1^r \dots \eta_n^r \bar{\Phi}_{m_1}\left(a_1 x_1+\frac{r}{k}, \eta_1\right) \dots \bar{\Phi}_{m_n}\left(a_n x_n+\frac{r}{k}, \eta_n\right),$$

where now

$$K = \prod_{i=1}^{n} \frac{1-\zeta_i}{1-\eta_i} c_i^{-m_i}.$$

Further specialization can be obtained by taking $e_1 = \dots = e_n$ and $f(x, \zeta) = \bar{B}_m(x)$, in which case (4. 2) must be used in place of (6. 2). We find that

$$(6.4)\qquad \sum_{r=0}^{Ak-1} \eta_1^r \dots \eta_n^r \bar{B}_{m_1}\left(x_1+\frac{r}{a_1 k}\right) \dots \bar{B}_{m_n}\left(x_n+\frac{r}{a_n k}\right) =$$

$$= K \sum_{r=0}^{k-1} \eta_1^r \dots \eta_n^r \bar{\Phi}_{m_1-1}\left(a_1 x_1+\frac{r}{k}, \eta_1\right) \dots \bar{\Phi}_{m_n-1}\left(a_n x_n+\frac{r}{k}, \eta_n\right),$$

where each $\eta_i \neq 1$ and

$$K = \prod_{i=1}^{n} \frac{m_i}{(\eta_i - 1) a_i^{m_i - 1}}.$$

If n (6. 4) some $\eta_i = 1$, the formula requires modification.

DUKE UNIVERSITY,
DURHAM, USA

(Received 16 June 1959)

References

[1] L. Carlitz, Some finite summation formulas of arithmetic character, *Publ. Math. Debrecen*, **6** (1959) (under press).

[2 M. Mikolás, Integral formulas of arithmetical characteristics relating to the zeta-function of Hurwitz, *Publ. Math. Debrecen*, **5** (1957), pp. 44—53.

[3] L. J. Mordell, Integral formulae of arithmetical character, *Journal London Math. Soc.*, **33** (1958), pp. 371—375.

NOTWENDIGE BEDINGUNGEN FÜR DIE DISKONTINUIERLICHEN LÖSUNGEN VON DEN VARIATIONSPROBLEMEN n-TER ORDNUNG

Von

A. KÓSA (Budapest)

(*Vorgelegt von* A. RÉNYI)

§ 1. Einleitende Bemerkungen und Formulierung des Problems

Bekanntlich besteht die einfachste Variationsaufgabe in der Bestimmung der Funktionen einer gegebenen — sog. zulässigen — Funktionenklasse,[1] welche Funktionen (wenn solche überhaupt existieren) für die Operation

$$I[y]=\int_{x_1}^{x_2} f(x,y,y')\,dx \tag{1.1}$$

ein Extremum liefern. Die Schwierigkeit und auch die Lösbarkeit des Problems hängt von der Auswahl der zulässigen Funktionenklasse ab. In der klassischen Theorie der Variationsrechnung spielen die nach den folgenden Gesichtspunkten ausgewählten Fälle eine ausgezeichnete Rolle:

a) Wenn wir die das Extremum liefernden Funktionen mit Hilfe der Auflösung einer Differentialgleichung bestimmen wollen, so müssen wir voraussetzen, daß die zulässigen Funktionon zweimal stetig differenzierbar sind.

b) Die engste Funktionenklasse, in welcher alle vorhandenen Funktionen noch stetig sind, ist die Klasse der einmal stetig differenzierbaren Funktionen.

c) Wegen ihrer mehrseitigen wichtigen Anwendungen ist es zweckmäßig, die Operation auch für die Klasse der sog. diskontinuierlichen Funktionen zu definieren. Darunter verstehen wir in unserem Falle Funktionen, welche stückweise stetig differenzierbar sind.[2] Die Untersuchung der Operation für Funktionen dieser Art bedeutet eine wesentliche Vereinfachung auch in der Berechnung der unter a) und b) beschriebenen Funktionen.

Wenden wir uns jetzt auf das sog. Variationsproblem n-ter Ordnung. Darunter verstehen wir ein Problem in der Form

$$I[y]=\int_{x_1}^{x_2} f(x,y,y',\ldots,y^{(n)})\,dx; \tag{1.2}$$

[1] Die dazugehörenden Funktionen werden zulässige Funktionen genannt.

[2] D. h. solche stetige Funktionen, deren Ableitung in endlich vielen Punkten endliche Sprünge haben darf.

die Operation hängt hier außer der Funktion auch noch von ihren ersten n Ableitungen ab.

Es erhebt sich die Frage, wie sich die obenerwähnten zulässigen Funktionenklassen im Falle des Problems n-ter Ordnung verändern. Bloß der Fall der diskontinuierlichen Funktionen kann sich als problematisch erweisen. Nämlich in Anbetracht der obenerwähnten Gesichtspunkte entsprechen für a) bzw. b) offensichtlich die $2n$-mal bzw. n-mal stetig differenzierbaren Funktionen.

Die Übertragung der diskontinuierlichen Funktionenklasse hat mehrere Möglichkeiten.

A. Behalten wir diejenige Eigenschaft der stückweise stetig differenzierbaren Funktionen, daß eine Unstetigkeit *nur in der letzten Ableitung*[3] auftreten kann, so gelangen wir zu solchen Funktionen, die mit ihren ersten $n-1$ Ableitungen stetig sind, die n-te Ableitung aber nur stückweise stetig ist.

B. Behalten wir aber diejenige Eigenschaft der stückweise stetig differenzierbaren Funktionen, daß *nur die Funktion selbst* überall stetig ist, so gelangen wir zu solchen Funktionen, die selbst stetig, ihre sämtlichen Ableitungen bis zur n-ten Ordnung inklusive aber nur stückweise stetig sind.[4]

In der Literatur finden wir — meines Wissens — nur Diskontinuierlichkeit von Typ **A**,[5] obwohl **B** eine ebenso naheliegende Verallgemeinerung von c) bedeutet wie **A**, die Diskontinuierlichkeit **A** ist sogar in **B** enthalten. Die Wichtigkeit von **B** wird außerdem noch durch folgende Gesichtspunkte hervorgehoben:

1. Viele einfache Operationen nehmen ein Extremum nur bei Funktionen von Diskontinuierlichkeit **B** an.

Beispiel 1. Es sei $n=2$, die Grundfunktion habe die Form

$$f=(4-y''^2)^2, \tag{1.3}$$

und die Randbedingungen seien

$$\begin{aligned} y(-1)&=0, \quad y(1)=0,\\ y'(-1)&=0, \quad y'(1)=-2. \end{aligned} \tag{1.4}$$

Die Operation

$$I[y]=\int_{-1}^{1}(4-y''^2)^2\,dx \tag{1.5}$$

[3] D. h. in der ersten.

[4] Es bestehen offenbar auch zwischenliegende Verallgemeinerungsmöglichkeiten.

[5] Siehe z. B. [7], [9], [10].

nimmt bei der Funktion

$$y(x)=\begin{cases}(x+1)^2 & \text{für } -1\leqq x\leqq 0,\\ -x^2+1 & \text{für } \quad 0\leqq x\leqq 1\end{cases} \tag{1.6}$$

wegen $f\geqq 0$, $I[y(x)]=0$ ein absolutes Minimum an, und hier sind

$$y'(-0)=2,\quad y'(+0)=0,$$
$$y''(-0)=2,\quad y''(+0)=-2.$$

Es ist leicht einzusehen, daß die Operation (1.5) den Wert 0 neben den Randbedingungen (1.4) nur bei solchen stetigen Funktionen annehmen kann, die — wie auch die Funktion (1.6) — eine Diskontinuität von Typ **B** besitzen.

BEISPIEL 2. Es sei

$$I[y]=\int_0^1\{(1-y'^2)^2+y''^2\}\,dx \tag{1.7}$$

neben den Randbedingungen

$$y(0)=0,\quad y(1)=0,$$
$$y'(0)=1,\quad y'(1)=-1.$$

Die Operation (1.7) nimmt bei der Funktion

$$y(x)=\begin{cases}x & \text{für } 0\leqq x\leqq\frac{1}{2},\\ 1-x & \text{für } \frac{1}{2}\leqq x\leqq 1\end{cases} \tag{1.8}$$

offensichtlich ein absolutes Minimum an.

Hier sind

$$y'\left(\frac{1}{2}-0\right)=1,\quad y'\left(\frac{1}{2}+0\right)=-1,$$

dagegen aber

$$y''(x-0)=y''(x+0)\quad \text{für}\quad 0<x<1.$$

Das Ergebnis ist nicht überraschend, da die Lösungsfunktionen in beiden Beispielen aus zwei solchen Funktionen stetig zusammengesetzt sind, die auch allein ein absolutes Minimum liefern; bei solcher Zusammensetzung ist aber im allgemeinen das Auftreten der Diskontinuierlichkeit **B** naturgemäß.

2. Die Natur mehrerer Aufgaben schließt nicht aus, daß die Operation auch für eine Diskontinuität von Typ **B** besitzende Funktionen erklärt wird. Darauf bezieht sich das

Beispiel 3. Es seien in der Ebene zwei Punkte P_1, P_2 angegeben, und legen wir in beiden Punkten eine Richtung fest. Wir betrachten alle Funktionen $y(x)$, deren Kurven von unten konvex bzw. konkav sind, die die zwei Punkte verbinden, und deren Normalen in den Punkten P_1, P_2 mit den gegebenen Richtungen zusammenfallen. Suchen wir diejenige von diesen Kurven aus, deren Radiusvektor einen minimalen Flächeninhalt streift.[6] Wenn die Kurve von unten konvex ist, dann suchen wir das Minimum, wenn konkav, dann das Maximum der Operation

$$I[y]=\int_{x_1}^{x_2}\frac{(1+y'^2)^2}{y''}\,dx. \tag{1.9}$$

Die Figuren stellen in der Reihenfolge diejenigen Fälle dar, wo im Punkte C nur die erste, dann nur die zweite und endlich beide Ableitungen einen Sprung haben.

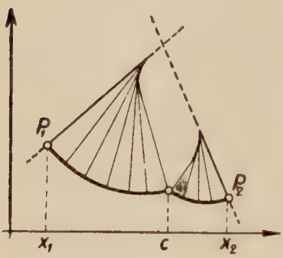


Fig. 1

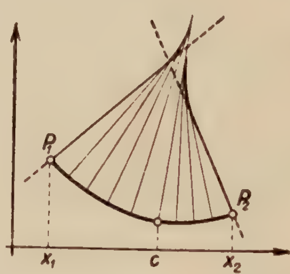


Fig. 2

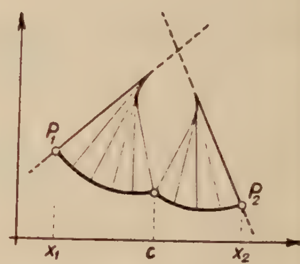


Fig. 3

Nach den Vorhergesagten betrachten wir das Problem n-ter Ordnung in folgender Form:

Sei G ein Gebiet[7] in der Ebene. Nehmen wir von der Grundfunktion

$$f(x, y, y', \ldots, y^{(n)}) \tag{1.10}$$

an, daß sie auf der Menge

$$R=\{(x,y)\in G,\ y^{(i)} \text{ beliebig endlich für } i=1,\ldots,n\}^{8} \tag{1.11}$$

alle partiellen Ableitungen bis zur 2-ten Ordnung inklusive besitzt, und alle diese stetige Funktionen ihrer sämtlichen $n+2$ Argumente sind.

[6] Siehe z. B. [9], S. 187.

[7] Darunter verstehen wir eine offene zusammenhängende Punktmenge.

[8] Alle unsere Ergebnisse bleiben gültig, wenn wir die Voraussetzung „$y^{(i)}$ beliebig endlich für $i=1,\ldots,n$" durch die folgende ersetzen: $a_i<y^{(i)}<b_i$, $-\infty\leq a_i<b_i\leq+\infty$, $i=1,\ldots,n$.

Wir nehmen zwei zu G gehörende Punkte (x_1, y_1), (x_2, y_2) $(x_1 \neq x_2)$, und geben $2(n-1)$ beliebige Zahlen $y_1^{(i)}, y_2^{(i)}$ $(i=1, \ldots, n-1)$ an.

Wir definieren jetzt die Funktionenklasse $D^{(n)}$ $(n=1, 2, \ldots)$ folgendermaßen:

$y(x) \in D^{(n)}$ im Intervall $[a, b]$, wenn $y(x)$ dort stetig ist, und das Intervall $[a, b]$ sich in eine endliche Anzahl von Teilintervallen zerlegen läßt derart, daß $y(x)$ im Inneren jedes Teilintervalls n-mal stetig differenzierbar ist, und jede Ableitung (bis zur n-ten Ordnung inklusive) in den Teilungspunkten endlichen rechts- und linksseitigen Grenzwert hat.[9]

Im folgenden werden wir irgendeinen Punkt x_0 des offenen Intervalls $a<x<b$ Eckenpunkt der Funktion $y(x) \in D^{(n)}$ nennen, wenn die Ungleichung

$$y^{(i)}(x_0-0) \neq y^{(i)}(x_0+0)$$

wenigstens für ein i $(i=1, \ldots, n)$ besteht.

Wir definieren nachher die zulässige Funktionenklasse M folgendermaßen: $y(x) \in M$, wenn

1. $y(x) \in D^{(n)}$ im Intervall $[x_1, x_2]$,
2. $y^{(i)}(x_1) = y_1^{(i)}$, $y^{(i)}(x_2) = y_2^{(i)}$ für $i=0, 1, \ldots, n-1$,[10]
3. $(x, y(x)) \in G$ für $x_1 \leqq x \leqq x_2$.

Für alle Funktionen $y(x) \in M$ hat die Operation

$$I[y(x)] = \int_{x_1}^{x_2} f(x, y(x), y'(x), \ldots, y^{(n)}(x))\,dx \tag{1.12}$$

einen bestimmten endlichen Wert.

Wir verstehen unter dem Abstand zweier zu M gehörenden Funktionen $y_1(x)$, $y_2(x)$ die Zahl

$$\varrho(y_1, y_2) = \max |y_1(x) - y_2(x)| \quad \text{für} \quad x_1 \leqq x \leqq x_2,$$

unter dem Abstand n-ter Ordnung aber die Zahl

$$\varrho_n(y_1, y_2) = \max \{\varrho(y_1, y_2), \varrho(y_1', y_2'), \ldots, \varrho(y_1^{(n)}, y_2^{(n)})\}.$$ [11]

Wir führen die folgende Definition ein:

Die Operation (1.12) nimmt bei der Funktion $y_0(x) \in M$ ein starkes relatives bzw. schwaches relatives Minimum an,[12] wenn eine positive Zahl

[9] In a soll nur der rechtsseitige, in b nur der linksseitige Grenzwert existieren.

[10] Hier und auch weiterhin verstehen wir unter $y^{(0)}(x)$ die Funktion $y(x)$ selbst: $y_1^{(0)} = y_1$, $y_2^{(0)} = y_2$.

[11] In den Teilungspunkten, wo irgendeine der Ableitungen nicht stetig ist, soll man die Grenzwerte der Funktionen von der entsprechenden Seite nehmen.

[12] Es gibt auch andere Erklärungsmöglichkeiten (siehe [8]).

ε existiert, so daß die Ungleichung

$$I[y(x)] \geqq I[y_0(x)]$$

für alle solche Funktionen $y(x) \in M$ gilt, für die $\varrho(y, y_0) < \varepsilon$ bzw. $\varrho_n(y, y_0) < \varepsilon$ besteht. (Das starke bzw. schwache relative Maximum wird ähnlicherweise definiert. Die zu beweisenden sämtlichen Sätze gelten auch für die Fälle des Maximums, wenn diese auf die Grundfunktion $-f(x, y, y', \ldots, y^{(n)})$ angewandt werden.)

Im weiteren geben wir notwendige Bedingungen für schwaches und starkes relatives Minimum an.

Im § 2 leiten wir als notwendige Bedingung ein neues System von Differentialgleichungen ab, verallgemeinern eine der sog. Weierstraß—Erdmannschen Eckenbedingungen, und teilen einige wichtige Folgen derselben mit. Im § 3 verallgemeinern wir die Legendresche notwendige Bedingung. Endlich im § 4 leiten wir die verallgemeinerte Weierstraßsche Bedingung ab, verallgemeinern die andere Weierstraß—Erdmannsche Eckenbedingung, und legen einige wichtige Folgen derselben fest.

§ 2. Die erste notwendige Bedingung: ein neues System von Differentialgleichungen. Eckenbedingung

Satz 1. *Wenn die Operation* (1.12) *bei der Funktion* $y(x) \in M$ *ein schwaches relatives Minimum annimmt, soll die Funktion* $y(x)$ *dem System von Differentialgleichungen*

$$f_{y^{(i)}}(x, y, y', \ldots, y^{(n)}) = 0 \quad \text{für} \quad i = 2, \ldots, n \tag{2.1}$$

genügen,[13] *außerdem soll die Relation*

$$\begin{aligned} &f_{y'}(x, y(x), y'(x-0), \ldots, y^{(n)}(x-0)) = \\ &= f_{y'}(x, y(x), y'(x+0), \ldots, y^{(n)}(x+0)) \end{aligned} \tag{2.2}$$

in jedem Punkte des Intervalls $x_1 < x < x_2$ *bestehen.*

Im Falle $n = 1$ fällt (2.2) mit der einen der sog. Weierstraß—Erdmannschen Eckenbedingungen zusammen.[14]

Beweis. Es sei c ein *beliebiger* Punkt des Intervalls $x_1 < x < x_2$. Wegen der Definition der Funktionenklasse $M \subset D^{(n)}$ kann eine solche positive Zahl δ angegeben werden, daß die Funktion $y(x)$ in den Teilintervallen

$$c - \delta < x < c, \quad c < x < c + \delta \tag{2.3}$$

[13] Man soll in den Eckenpunkten beiderseitige Grenzwerte in Betracht nehmen.

[14] Siehe z. B. [3], S. 143.

des Intervalls $x_1 < x < x_2$ keinen Eckenpunkt hat. Es seien a und b zwei solche Zahlen, für die

$$c-\delta < a < c, \quad c < b < c+\delta \tag{2.4}$$

sind, und nehmen wir eine beliebige, im Intervall $[a, b]$ der Funktionenklasse $D^{(n)}$ gehörende Funktion $\omega(x)$, welche in den Punkten a und b verschwindet und in den Intervallen

$$a < x < c, \quad c < x < b \tag{2.5}$$

keinen Eckenpunkt hat.

Betrachten wir nachher die Funktion

$$\eta(x) = \begin{cases} \omega(x) & \text{für} \quad a \leqq x \leqq b, \\ 0 & \text{für} \quad \begin{cases} x_1 \leqq x \leqq a, \\ b \leqq x \leqq x_2. \end{cases} \end{cases} \tag{2.6}$$

Es ist offenbar, daß die Funktionen

$$y(x) + \varepsilon\eta(x)$$

für hinreichend kleines $|\varepsilon| < \varepsilon_0$ $(\varepsilon_0 > 0)$ der Funktionenklasse M gehören. Die Funktion

$$v(\varepsilon) = I[y(x) + \varepsilon\eta(x)] \tag{2.7}$$

nimmt also auf Grund unserer Annahme bei $\varepsilon = 0$ ein relatives Minimum an. Ebenso ist (2. 7), wenn $|\varepsilon| < \varepsilon_0$ gilt, wegen der Voraussetzung differenzierbar, und die Derivation ist unter dem Integralzeichen durchzuführen.[15] Es soll also die Relation

$$v'(0) = 0 \tag{2.8}$$

bestehen.

Nach Differentiation erhalten wir unter Berücksichtigung von (2. 6)

$$v'(0) = \int_a^b \{\bar{f}_y \cdot \eta + \bar{f}_{y'} \cdot \eta' + \cdots + \bar{f}_{y^{(n)}} \cdot \eta^{(n)}\}\, dx = 0.^{16} \tag{2.9}$$

Es sei nun p irgendeine von den Zahlen $1, \ldots, n$, und untersuchen wir die Integrale

$$v_p[\eta] = \int_a^b \{\bar{f}_y \cdot \eta + \bar{f}_{y'} \cdot \eta' + \cdots + \bar{f}_{y^{(p)}} \cdot \eta^{(p)}\}\, dx. \tag{2.10}$$

Durch teilweise Integration des $(i+1)$-ten Gliedes des Integrands be-

[15] Siehe z. B. [2], S. 63.
[16] Wir bezeichnen $\bar{f}_{y^{(i)}} = f_{y^{(i)}}(x, y(x), y'(x), \ldots, y^{(n)}(x))$ für $i = 0, 1, \ldots, n$.

kommen wir:

$$(2.11)\quad \int_a^b \bar{f}_{y^{(i)}} \cdot \eta^{(i)}(x)\,dx = \Big[\int_c^x \bar{f}_{y^{(i)}}dx \cdot \eta^{(i)}(x)\Big]_{x=a}^{x=c-0} + \Big[\int_c^x \bar{f}_{y^{(i)}}dx \cdot \eta^{(i)}(x)\Big]_{x=c+0}^{x=b} - \\ - \int_a^c \Big(\int_c^x \bar{f}_{y^{(i)}}dx\Big)\cdot \eta^{(i+1)}(x)\,dx - \int_c^b \Big(\int_c^x \bar{f}_{y^{(i)}}dx\Big)\cdot \eta^{(i+1)}(x)\,dx.$$

Integrieren wir das auf der rechten Seite entstandene Integral wieder teilweise, so gewinnen wir nach $p-i$ Schritten

$$(2.12)\quad \int_a^b \bar{f}_{y^{(i)}} \cdot \eta^{(i)}(x)\,dx = \sum_{\alpha=1}^{p-i}(-1)^{\alpha-1}\Big\{\int_a^c\int_c^{x_\alpha}\cdots\int_c^{x_2}\bar{f}_{y^{(i)}}(x_1)\,dx_1\cdots dx_\alpha\cdot\eta^{(i+\alpha-1)}(a+0)+ \\ +\int_c^b\int_c^{x_\alpha}\cdots\int_c^{x_2}\bar{f}_{y^{(i)}}(x_1)\,dx_1\cdots dx_\alpha\cdot\eta^{(i+\alpha-1)}(b-0)\Big\}+ \\ +(-1)^{p-i}\Big\{\int_a^c\Big(\int_c^{x_{p-i+1}}\cdots\int_c^{x_2}\bar{f}_{y^{(i)}}(x_1)\,dx_1\cdots dx_{p-i}\Big)\cdot\eta^{(p)}(x_{p-i+1})\,dx_{p-i+1}+ \\ +\int_c^b\Big(\int_c^{x_{p-i+1}}\cdots\int_c^{x_2}\bar{f}_{y^{(i)}}(x_1)\,dx_1\cdots dx_{p-i}\Big)\cdot\eta^{(p)}(x_{p-i+1})\,dx_{p-i+1}\Big\}.$$

Führen wir die folgenden Bezeichnungen ein:

$$(2.13)\quad u_{\alpha,i}(x_\alpha) = \int_c^{x_\alpha}\cdots\int_c^{x_2}\bar{f}_{y^{(i)}}(x_1)\,dx_1\cdots dx_{\alpha-1} \\ \text{für}\quad i=0,1,\dots,p-1\quad(\alpha=1,2,\dots).$$

Nach dem Einsetzen der Funktionen (2.13) in (2.12), hat (2.10) die Gestalt

$$(2.14)\quad v_p[\eta] = \int_a^c\Big\{\sum_{i=0}^{p-1}(-1)^{p-i}u_{p-i+1,i}(x)+\bar{f}_{y^{(p)}}(x)\Big\}\cdot\eta^{(p)}(x)\,dx+ \\ +\int_c^b\Big\{\sum_{i=0}^{p-1}(-1)^{p-i}u_{p-i+1,i}(x)+\bar{f}_{y^{(p)}}(x)\Big\}\cdot\eta^{(p)}(x)\,dx+ \\ +\sum_{i=0}^{p-1}\sum_{\alpha=1}^{p-i}(-1)^{\alpha-1}\Big\{\int_a^c u_{\alpha,i}(x)\,dx\cdot\eta^{(i+\alpha-1)}(a+0)+ \\ +\int_c^b u_{\alpha,i}(x)\,dx\cdot\eta^{(i+\alpha-1)}(b-0)\Big\}\quad \text{für}\quad p=1,\dots,n.$$

Wir betrachten jetzt die Funktionenfolgen

$$(2.15)\qquad \varphi_{p,k}(x)=\begin{cases} \dfrac{k}{p!}\left(x-c+\dfrac{1}{k}\right)^{p} & \text{für}\quad c-\dfrac{1}{k}\leqq x\leqq c,\\ -\dfrac{1}{p!}\left(\dfrac{1}{k}\right)^{p-2}\left(x-c-\dfrac{1}{k}\right) & \text{für}\quad c\leqq x\leqq c+\dfrac{1}{k},\\ 0 & \text{für}\quad \begin{cases} x_1\leqq x\leqq c-\dfrac{1}{k},\\ c+\dfrac{1}{k}\leqq x\leqq x_2\end{cases}\end{cases}$$

und

$$(2.16)\qquad \psi_{p,k}(x)=\begin{cases} \dfrac{1}{p!}\left(\dfrac{1}{k}\right)^{p-2}\left(x-c+\dfrac{1}{k}\right) & \text{für}\quad c-\dfrac{1}{k}\leqq x\leqq c,\\ \dfrac{k}{p!}(-1)^{p}\left(x-c-\dfrac{1}{k}\right)^{p} & \text{für}\quad c\leqq x\leqq c+\dfrac{1}{k},\\ 0 & \text{für}\quad \begin{cases} x_1\leqq x\leqq c-\dfrac{1}{k},\\ c+\dfrac{1}{k}\leqq x\leqq x_2\end{cases}\end{cases}$$

$$\text{für}\quad p=1,\ldots,n;\quad k=k_0+1,\,k_0+2,\ldots,$$

wo k_0 eine solche positive ganze Zahl ist, für die $\frac{1}{k_0}<\delta$ gilt.

Es ist klar, daß die Funktionen (2.15) und (2.16) alle Bedingungen erfüllen, welche wir von der Funktion $\eta(x)$ vorausgesetzt haben; hier nehmen die Rolle der Zahlen a, b die Zahlen $c-\frac{1}{k}$, $c+\frac{1}{k}$ über.

Wir erhalten nach einer einfachen Rechnung für alle $k=k_0+1, k_0+2,\ldots$ folgendes:

$$(2.17)\qquad \varphi_{1,k}(x)=\psi_{1,k}(x)\quad \text{für}\quad x_1\leqq x\leqq x_2,$$

$$(2.18)\qquad \varphi'_{1,k}(x)=\begin{cases} k & \text{für}\quad c-\dfrac{1}{k}<x<c,\\ -k & \text{für}\quad c<x<c+\dfrac{1}{k},\end{cases}$$

$$(2.19)\qquad \varphi^{(i)}_{p,k}\left(c-\frac{1}{k}\right)=0\quad \text{für}\quad \begin{cases} p=2,\ldots,n,\\ i=0,1,\ldots,p-1,\end{cases}$$

$$(2.20)\qquad \psi^{(i)}_{p,k}\left(c+\frac{1}{k}\right)=0\quad \text{für}\quad \begin{cases} p=2,\ldots,n,\\ i=0,1,\ldots,p-1,\end{cases}$$

$$(2.21)\qquad \varphi^{(p)}_{p,k}(x)=\begin{cases} k & \text{für}\quad c-\frac{1}{k}<x<c,\\ 0 & \text{für}\quad c<x<c+\frac{1}{k}\end{cases}$$

$$(p=2,\ldots,n);$$

$$(2.22)\qquad \psi^{(p)}_{p,k}(x)=\begin{cases} 0 & \text{für}\quad c-\frac{1}{k}<x<c,\\ (-1)^p k & \text{für}\quad c<x<c+\frac{1}{k}\end{cases}$$

$$(p=2,\ldots,n);$$

$$(2.23)\qquad \varphi^{(\alpha)}_{p,k}(x-0)=\varphi^{(\alpha)}_{p,k}(x+0)=\psi^{(\alpha)}_{p,k}(x-0)=\psi^{(\alpha)}_{p,k}(x+0)=0$$

$$\text{für}\quad x_1<x<x_2;\quad p=1,\ldots,n;\quad \alpha=p+1,\ldots,n;$$

$$(2.24)\qquad \begin{gathered} |\varphi^{(i)}_{p,k}(x-0)|\leqq 1,\quad |\varphi^{(i)}_{p,k}(x+0)|\leqq 1;\\ |\psi^{(i)}_{p,k}(x-0)|\leqq 1,\quad |\psi^{(i)}_{p,k}(x+0)|\leqq 1 \end{gathered}$$

$$\text{für}\quad x_1<x<x_2;\quad p=1,\ldots,n;\quad i=0,1,\ldots,p-1.$$

Nach dem Einsetzen der Funktionen (2.15) und (2.16) in (2.10), bekommen wir wegen (2.23)

$$(2.25)\qquad v_p[\varphi_{p,k}]=0,\quad v_p[\psi_{p,k}]=0\quad \text{für}\quad p=1,\ldots,n;\quad k=k_0+1,k_0+2,\ldots.$$

Zuerst sei $p=1$. Die entsprechende Gleichheit (2.25) hat unter Benützung von (2.14) wegen (2.18) und (2.23) die Form

$$(2.26)\qquad \begin{aligned} v_1[\varphi_{1,k}]=\int_{c-\frac{1}{k}}^{c}\{-u_{2,0}(x)+\bar f_{y'}(x)\}k\,dx+\\ +\int_{c}^{c+\frac{1}{k}}\{-u_{2,0}(x)+\bar f_{y'}(x)\}(-k)\,dx=0. \end{aligned}$$

Wenden wir jetzt den Mittelwertsatz auf die Integrale in (2.26) an. Durch eine Umordnung erhalten wir:

$$(2.27)\qquad \bar f_{y'}(\xi_k)-\bar f_{y'}(\xi_k^*)=u_{2,0}(\xi_k)-u_{2,0}(\xi_k^*),$$

wo

$$(2.28)\qquad c-\frac{1}{k}<\xi_k<c,\quad c<\xi_k^*<c+\frac{1}{k}\qquad (k=k_0+1,k_0+2,\ldots)$$

sind.

Wegen der Definition der Funktionen (2.13) und wegen der Ungleichungen (2.28) streben aber beide Glieder auf der rechten Seite von (2.27) für $k \to +\infty$ gegen 0. Es wird also

$$\bar{f}_{y'}(c-0) = \bar{f}_{y'}(c+0). \tag{2.29}$$

Die Relation (2.2) ist damit bewiesen.

Um das System (2.1) zu beweisen, nehmen wir die Gleichheiten (2.25) für $p \geqq 2$. Wir erhalten aus (2.14) unter Berücksichtigung von (2.19), (2.20), (2.21) und (2.22) folgendes:

$$\begin{aligned} v_p[\varphi_{p,k}] = & \int_{c-\frac{1}{k}}^{c} \left\{ \sum_{i=0}^{p-1} (-1)^{p-i} u_{p-i+1,i}(x) + \bar{f}_{y^{(p)}}(x) \right\} k\,dx + \\ & + \int_{c}^{c+\frac{1}{k}} \{u_{1,1}(x) - u_{2,0}(x)\}\,dx \cdot \varphi'_{p,k}\left(c + \frac{1}{k} - 0\right) \end{aligned} \tag{2.30}$$

und

$$\begin{aligned} v_p[\psi_{p,k}] = & \int_{c}^{c+\frac{1}{k}} \left\{ \sum_{i=0}^{p-1} (-1)^{p-i} u_{p-i+1,i}(x) + \bar{f}_{y^{(p)}}(x) \right\} (-1)^p k\,dx + \\ & + \int_{c-\frac{1}{k}}^{c} \{u_{1,1}(x) - u_{2,0}(x)\}\,dx \cdot \psi'_{p,k}\left(c - \frac{1}{k} + 0\right) \end{aligned} \tag{2.31}$$

$$\text{für} \quad p = 2, \ldots, n; \quad k = k_0 + 1, k_0 + 2, \ldots.$$

Wenden wir auch jetzt den Mittelwertsatz auf das erste Integral in (2.30) und (2.31) an. Wir gewinnen nach einer Umordnung

$$\begin{aligned} \bar{f}_{y^{(p)}}(\tau_k) = & \sum_{i=0}^{p-1} (-1)^{p-i+1} u_{p-i+1,i}(\tau_k) + \\ & + \int_{c}^{c+\frac{1}{k}} \{u_{2,0}(x) - u_{1,1}(x)\}\,dx \cdot \varphi'_{p,k}\left(c + \frac{1}{k} - 0\right) \end{aligned} \tag{2.32}$$

und

$$(2.33)\qquad \bar{f}_{y^{(p)}}(\vartheta_k)=\sum_{i=0}^{p-1}(-1)^{2p-i+1}u_{p-i+1,i}(\vartheta_k)+$$
$$+(-1)^p\int_{c-\frac{1}{k}}^{c}\{u_{2,0}(x)-u_{1,1}(x)\}\,dx\cdot\psi'_{p,k}\left(c-\frac{1}{k}+0\right)\quad\text{für}\quad p=2,\ldots,n,$$

wo

$$(2.34)\qquad c-\frac{1}{k}<\tau_k<c,\quad c<\vartheta_k<c+\frac{1}{k}\qquad (k=k_0+1,k_0+2,\ldots)$$

sind.

Auf der rechten Seite beider Gleichheiten (2.32) und (2.33) stehen endlich viele Glieder. Diese streben aber wegen der Definition der Funktionen (2.13) und wegen der Ungleichungen (2.24) und (2.34) für $k\to+\infty$ gegen 0. Wir erhielten also

$$\bar{f}_{y^{(p)}}(c-0)=0,\quad \bar{f}_{y^{(p)}}(c+0)=0\quad\text{für}\quad p=2,\ldots,n.$$

Unser Satz ist damit bewiesen.[17]

Folgen und Bemerkungen.

1. Im Beispiel 1 ist $f_{y'}\equiv 0$, und auch im Beispiel 2 gilt

$$f_{y'}(x,y(x),y'(x),y''(x))=0\quad\text{für}\quad 0\leqq x\leqq 1,$$

wo $y(x)$ die das Minimum liefernde Funktion (1.8) bedeutet. Daraus und aus dem Bestehen des Systems (2.1) kann man daran denken, daß die Relation (2.2) immer so erfüllt wird, daß der gemeinsame Grenzwert in jedem Punkt des Intervalls $x_1<x<x_2$ 0 ist; dies erfolgt aber aus der obigen Beweisführung nicht. Darauf, daß es im allgemeinen nicht so ist, d. h., daß wir in dieser Beziehung über die Funktion $f_{y'}$ nicht mehr behaupten können als die Relation (2.2), bezieht sich

Beispiel 4. Wir betrachten die Operation

$$(2.35)\qquad I[y]=\int_0^1\{(y'-2)^2+y''^2\}\,dx$$

neben den Randbedingungen

$$(2.36)\qquad \begin{aligned} y(0)&=0, & y(1)&=1,\\ y'(0)&=1, & y'(1)&=1, \end{aligned}$$

[17] Unsere obige Beweisführung bleibt auch im Falle gültig, wenn die Grundfunktion $f(x,y,y',\ldots,y^{(n)})$ bei endlich vielen x nicht stetig ist, die Funktionen $\bar{f}_{y^{(i)}}(x)$ $(i=1,\ldots,n)$ aber stückweise stetig sind.

und die Funktion

(2.37) $$y_0(x) = x \quad \text{für} \quad 0 \leqq x \leqq 1.$$

Hier sind

(2.38) $$y_0'(x) = 1, \quad y_0''(x) = 0 \quad \text{für} \quad 0 < x < 1.$$

Schreiben wir jetzt die zulässigen Funktionen in der Form

(2.39) $$y(x) = y_0(x) + \eta(x).$$

Das bedeutet, daß die Funktionen $\eta(x)$ der Funktionenklasse D'' gehören und außerdem die Bedingungen

(2.40) $$\eta(0) = \eta(1) = \eta'(0) = \eta'(1) = 0$$

erfüllen sollen, übrigens aber sie beliebig sein können.

Für die totale Variation der Operation (2.35) bekommen wir unter Berücksichtigung von (2.38) folgendes:

(2.41) $$\begin{aligned} &I[y(x)] - I[y_0(x)] = \\ &= \int_0^1 \{(y_0'(x) + \eta'(x) - 2)^2 + (y_0''(x) + \eta''(x))^2 - (y_0'(x) - 2)^2 - y_0''(x)^2\}\,dx = \\ &= \int_0^1 \{-2\eta'(x) + \eta'^2(x) + \eta''^2(x)\}\,dx. \end{aligned}$$

Da wegen (2.40)

(2.42) $$\int_0^1 \eta'(x)\,dx = 0$$

ist, erhalten wir aus (2.41)

(2.43) $$I[y(x)] - I[y_0(x)] = \int_0^1 (\eta'^2(x) + \eta''^2(x))\,dx \geqq 0.$$

Die Ungleichung (2.43) bedeutet aber, daß die Operation (2.35) bei der Funktion (2.37) neben den Randbedingungen (2.36) ein absolutes Minimum[18] annimmt. Es gilt aber für die Funktion $y_0(x)$ die Relation

(2.44) $$f_{y'}(x, y_0(x), y_0'(x), y_0''(x)) = 2(y_0'(x) - 2) = -2 \neq 0 \quad \text{für} \quad 0 \leqq x \leqq 1.$$

Bei diesem Beispiel ist es bemerkenswert, daß, wenn wir andere Randbedingungen nehmen:

(2.45) $$\begin{aligned} y(a) &= a_0, \quad y(b) = b_0, \\ y'(a) &= a_1, \quad y'(b) = b_1, \end{aligned}$$

[18] In diesem Falle natürlich auch ein schwaches relatives Minimum.

es keine den Randbedingungen (2. 45) genügende Funktion gibt — den einzigen Fall

$$(2.46) \qquad a_1 = b_1 = \frac{b_0 - a_0}{b - a}$$

ausgenommen —, welche die Gleichungen (2. 1) und (2. 2) erfüllt. Die Lösung der Gleichung $f_{y''} = 2y'' = 0$ ist hier nämlich die Schar der Geraden $y = \alpha x + \beta$, die zwei Punkte sind also — den Fall (2. 46) ausgenommen — nur mit gebrochener Linie zu verbinden, und in den Eckenpunkten kann die Bedingung (2. 2) wegen $f_{y'} = 2(y' - 2)$ nicht erfüllt werden.

2. Nehmen wir an, daß die das schwache relative Minimum liefernde Funktion $y_0(x) \in M$ im Intervall $a < x < b$ $(x_1 \leqq a < b \leqq x_2)$ keinen Eckenpunkt hat. Untersuchen wir jetzt die Operation

$$(2.47) \qquad I^*[y] = \int_a^b f(x, y, y', \ldots, y^{(n)})\,dx$$

neben den Randbedingungen

$$(2.48) \quad y^{(i)}(a) = y_0^{(i)}(a+0), \quad y^{(i)}(b) = y_0^{(i)}(b-0) \quad \text{für} \quad i = 0, 1, \ldots, n-1.$$

Es ist offenbar, daß die Operation (2. 47) bei der Funktion $y_0(x)$ in Bezug auf alle solche n-mal stetig differenzierbaren Funktionen ein schwaches relatives Minimum annimmt, die den Randbedingungen (2. 48) genügen, und für die im Intervall $a < x < b$ $\varrho_n(y, y_0) < \varepsilon$ gilt, wo ε eine hinreichend kleine positive Zahl ist. In diesem Falle soll aber die Funktion $y_0(x)$ — wie bekannt — die sog. Euler—Poissonsche Gleichung

$$(2.49) \qquad f_y - \frac{d}{dx} f_{y'} + \frac{d^2}{dx^2} f_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n} f_{y^{(n)}} = 0$$

im Intervall $a < x < b$ erfüllen.[19] Wegen (2. 1) ist aber jedes Glied in (2. 49) vom dritten Glied ausgehend gleich 0. Damit bewiesen wir den

SATZ 2. *Wenn die Operation* (1. 12) *bei der Funktion* $y(x) \in M$ *ein schwaches relatives Minimum annimmt, soll die Funktion* $y(x)$ *der Gleichung*

$$(2.50) \qquad f_y(x, y, y', \ldots, y^{(n)}) - \frac{d}{dx} f_{y'}(x, y, y', \ldots, y^{(n)}) = 0$$

im Intervall $x_1 < x < x_2$ *— die Eckenpunkte ausgenommen — genügen.*[20]

[19] Siehe z. B. [1], S. 205.

[20] Im Zusammenhang mit einem anderen Problem hatte auch H. H. PIXLEY im Falle $n = 2$ die entsprechenden Gleichungen $f_{y''} = 0$, $f_y - \frac{d}{dx} f_{y'} = 0$ erhalten. Er hatte die Operation für solche Funktionen erklärt, bei denen in endlich vielen Punkten nicht nur die Ableitungen, sondern *auch selbst die Funktionen* eine Unstetigkeit besitzen können (siehe [6]).

3. Die Möglichkeit, welche für die Aufsuchung der das Minimum liefernden Funktionen durch das System (2. 1) gegeben ist, erweist sich wesentlich einfacher als die, welche aus der Euler—Poissonschen Gleichung folgt, und die Gleichungen (2. 1) richten sich auch besser der Natur der Aufgabe. Die Euler—Poissonsche Gleichung kann wesentlich nur dann gebraucht werden, wenn wir voraussetzen, daß die das Minimum liefernden Funktionen wenigstens $2n$-mal stetig differenzierbar sind; in diesem Falle geht die Euler—Poissonsche Gleichung in eine gewöhnliche Differentialgleichung $2n$-ter Ordnung über. Demgegenüber sind die Gleichungen des Systems (2. 1) gewöhnliche Differentialgleichungen n-ter Ordnung. Auch die Gleichung (2. 50) ist vom einfacheren Typ als die Euler—Poissonsche Gleichung, da, wenn die das Minimum liefernden Funktionen wenigstens $(n+1)$-mal stetig differenzierbar sind, so geht (2. 50) in eine gewöhnliche Differentialgleichung $(n+1)$-ter Ordnung über.

Es werden noch folgende betont: Die Gleichung (2. 50) und die Eckenbedingung (2. 2) sind für $n=1$ bekannt.[21] Die obigen Resultate sagen aus, daß diese — im Falle $n=1$ bekannten — Gleichungen *ihre Gestalt in den Fällen* $n \geqq 2$ *nicht verändern; die Abweichung besteht nur in der Verschiedenheit der Grundfunktionen.* Für die nichtdiskontinuierlichen Probleme[22] ist eine scharf gegenüberstehende Tatsache wahr: die Rolle der Euler—Lagrangeschen Gleichung wird durch die auch in der Form kompliziertere Euler—Poissonsche Gleichung übernommen.[23]

4. Die Aufsuchung der das Minimum liefernden Funktionen ist nun im allgemeinen folgendermaßen zweckmäßig: Wir lösen von den Gleichungen des Systems (2. 1) die einfachste auf. Nachher entscheiden wir durch Einsetzen, ob es unter diesen Lösungen auch solche gibt, die auch den weiteren Gleichungen des Systems (2. 1) und der Gleichung (2. 50) genügen. Wenn es solche gibt, so versuchen wir von diesen — unter Berücksichtigung der Eckenbedingung (2. 1) — die Konstruktion der die gegebenen Randbedingungen erfüllenden Funktionen.

Es ist schon im voraus zu sehen, daß — da die gesuchten Funktionen mehreren Gleichungen genügen sollen — die Randbedingungen im allgemeinen nicht beliebigerweise vorgeschrieben werden können.[24] Wir können im allgemeinen in beiden Endpunkten so viele Randbedingungen angeben, wieviel

[21] Im Falle $n=1$ ist die Gleichung (2. 50) die bekannte Euler—Lagrangesche Gleichung. Siehe z. B. [2], S. 30.

[22] Darunter verstehen wir das Problem, wo auch noch die letzte vorhandene Ableitung stetig ist.

[23] Siehe z. B. [1], S. 205.

[24] Siehe die Definition der zulässigen Funktionenklasse, S. 27.

Konstanten die allgemeine gemeinsame Lösung der Gleichungen (2.1) und (2.50) besitzt. Betrachten wir dazu die folgenden Beispiele:

BEISPIEL 5. Sei $f=f(y^{(n)})$ $(n\geqq 2)$, wo die Funktion $f'(u)$ wenigstens eine Nullstelle hat. Die Gleichungen (2.1) und (2.50) reduzieren sich jetzt auf eine einzige Gleichung

$$f'(y^{(n)})=0. \tag{2.51}$$

Die allgemeine Lösung der Gleichung (2.51) ist

$$y=a_n x^n+a_{n-1}x^{n-1}+\cdots+a_1 x+a_0, \tag{2.52}$$

wo die Zahlen $a_0, a_1, \ldots, a_{n-1}$ beliebig sind, a_n aber eine solche Zahl ist, für die $f'(n!\,a_n)=0$ gilt. In diesem Falle können in beiden Endpunkten Randbedingungen der Anzahl n vorgeschrieben werden. Die Eckenbedingung (2.2) erfüllt sich offensichtlich.

BEISPIEL 6. Sei

$$f=y^2(1-y')^2(1+y''^2)+y'''^2. \tag{2.53}$$

Die Gleichungen (2.1) lauten hier:

$$f_{y''}=2y^2(1-y')^2y''=0, \tag{2.54}$$

$$f_{y'''}=2y'''=0 \tag{2.55}$$

Die allgemeine Lösung von (2.54) ist die Geradenschar

$$y=ax+b. \tag{2.56}$$

Alle Funktionen (2.56) genügen offensichtlich der Gleichung (2.55).

Wir bekommen nach einer einfachen Rechnung, daß die entsprechende Gleichung

$$f_y-\frac{d}{dx}f_{y'}=2y(1-y')^2(1+y''^2)+\frac{d}{dx}\{2y^2(1-y')(1+y''^2)\}=0 \tag{2.57}$$

von den Funktionen (2.56) nur durch die Funktionen

$$\begin{cases} y=0, \\ y=\pm x+b \end{cases} \tag{2.58}$$

erfüllt wird, wo b eine beliebige Konstante ist.

Aus der Gestalt der Funktionen (2.58) geht hervor, daß wir in den Endpunkten x_1, x_2 nur die Funktionenwerte vorschreiben können, und auch

diese nicht völlig beliebig, sondern so, daß eine der Bedingungen

$$|y(x_1)|+|y(x_2)|\leqq x_2-x_1, \qquad (2.59)$$
$$|y(x_2)-y(x_1)|=x_2-x_1$$

bestehen muß.

Da

$$f_{y'}\doteq-2y^2(1-y')(1+y'^2), \qquad (2.60)$$

ist es leicht einzusehen, daß entweder die Funktion

$$y(x)=\begin{cases} y(x_1)-\operatorname{sg} y(x_1)(x-x_1) & \text{für } x_1\leqq x\leqq x_1+|y(x_1)|,\\ 0 & \text{für } x_1+|y(x_1)|\leqq x\leqq x_2-|y(x_2)|,\\ y(x_2)+\operatorname{sg} y(x_2)(x-x_2) & \text{für } x_2-|y(x_2)|\leqq x\leqq x_2\end{cases} \qquad (2.61)$$

oder eine der Funktionen (2.58) im Falle der Erfüllung der Voraussetzung (2.59) die Punkte $(x_1, y(x_1))$ und $(x_2, y(x_2))$ verbindet, und die Eckenbedingung in den zufälligen Eckenpunkten erfüllt. Aus (2.60) ist auch klar, daß die Eckenbedingung (2.2) in solchen Eckenpunkten der aus den Funktionen (2.58) stetig zusammengesetzten Funktionen, in denen $y(x)\neq 0$ ist, nicht erfüllen kann. Wir erhielten also, daß es im Falle der Erfüllung der Voraussetzung (2.59) gerade eine solche Funktion gibt, die sämtliche bisher erhaltene notwendige Bedingungen erfüllt, und deren Kurve die Punkte $(x_1, y(x_1))$ und $(x_2, y(x_2))$ verbindet.

Beispiel 7. Sei

$$f=\frac{1}{2}(y'^2+y''^2)+e^y. \qquad (2.62)$$

Die entsprechenden Gleichungen sind

$$f_{y''}=y''=0 \qquad (2.63)$$

und

$$f_y-\frac{d}{dx}f_{y'}=e^y-\frac{d}{dx}y'=e^y-y''=0. \qquad (2.64)$$

Die Gleichungen (2.63) und (2.64) haben offenbar keine gemeinsame Lösung.

Endlich betrachten wir das Beispiel 3. Dort war

$$f=\frac{(1+y'^2)^2}{y''}.$$

Es ist klar, daß es keine solche Funktion gibt, die der Gleichung

$$f_{y''} = -\frac{(1+y'^2)^2}{y''^2} = 0 \tag{2.65}$$

genügt. Das im Beispiel 3 aufgeworfene Problem hat also keine Lösung.[25]

5. Auf die Unstetigkeit in den Ableitungen bezieht sich in gewisser Hinsicht der

SATZ 3. *Sei $y(x) \in M$ und setzen wir voraus, daß in irgendwelchem Punkt x_0 des Intervalls $x_1 < x < x_2$ für ein gewisses l $(1 \leqq l \leqq n)$*

$$y^{(l)}(x_0 - 0) \neq y^{(l)}(x_0 + 0) \tag{2.66}$$

gilt, sonst aber

$$y^{(i)}(x_0 - 0) = y^{(i)}(x_0 + 0) \tag{2.67}$$

ist für $i = 1, \ldots, n$; $i \neq l$; außerdem sei

$$f_{y^{(l)}y^{(l)}}(x_0, y(x_0), \ldots, y^{(l-1)}(x_0), \tau, y^{(l+1)}(x_0), \ldots) \neq 0 \quad \text{für} \quad a \leqq \tau \leqq b, \tag{2.68}$$

wo

$$a = \min\{y^{(l)}(x_0 - 0), y^{(l)}(x_0 + 0)\}, \quad b = \max\{y^{(l)}(x_0 - 0), y^{(l)}(x_0 + 0)\}$$

sind. In diesem Falle kann die Operation (1.12) bei $y(x)$ auch noch kein schwaches relatives Extremum annehmen.[26]

Die Beweisführung ist sehr einfach. Die Funktion

$$\psi(\tau) = f_{y^{(l)}}(x_0, y(x_0), \ldots, y^{(l-1)}(x_0), \tau, y^{(l+1)}(x_0), \ldots) \tag{2.69}$$

ist wegen (2.68) eine streng monotone Funktion der Veränderlichen τ im Intervall $[a, b]$, also

$$\begin{aligned} &f_{y^{(l)}}(x_0, y(x_0), y'(x_0 - 0), \ldots, y^{(n)}(x_0 - 0)) \neq \\ &\neq f_{y^{(l)}}(x_0, y(x_0), y'(x_0 + 0), \ldots, y^{(n)}(x_0 + 0)), \end{aligned} \tag{2.70}$$

d. h. die Funktion $y(x)$ genügt wegen (2.70) entweder den Gleichungen (2.1) oder der Eckenbedingung (2.2) nicht. Die Behauptung ist also bewiesen.

[25] Diese Tatsache ist aber auch auf elementarem Weg leicht einzusehen. Für *zweimal stetig differenzierbare Funktionen* siehe z. B. [9], S. 187.

[26] Für $n = 1$ ist der Satz bekannt. Siehe [5], § 4, und [3], S. 143.

§ 3. Verallgemeinerung der Legendreschen Bedingung

SATZ 4. *Wenn die Operation* (1.12) *bei der Funktion* $y(x) \in M$ *ein schwaches relatives Minimum annimmt, sollen die Ungleichungen*

$$(3.1) \qquad \bar{f}_{y^{(i)}y^{(i)}}(x) = f_{y^{(i)}y^{(i)}}(x, y(x), y'(x), \ldots, y^{(n)}(x)) \geqq 0 \quad \textit{für} \quad i = 1, \ldots, n$$

im ganzen Intervall $[x_1, x_2]$ *bestehen.*[27]

Wenn $y'(x), \ldots, y^{(n-1)}(x)$ überall stetig sind, so fällt (3.1) im Falle $i = n$ mit der bekannten Legendreschen Bedingung zusammen.[28]

Es genügt unsere Behauptung nur für solche inneren Punkte des Intervalls $[x_1, x_2]$ nachzuweisen, in denen die Ableitungen $y'(x), \ldots, y^{(n)}(x)$ stetig sind; die Behauptung in den übrigen Punkten folgt dann nämlich aus der Definition der Funktionenklasse $M \subset D^{(n)}$.

Wir führen den Beweis indirekt für die p-te von den Ungleichungen (3.1); p kann die Werte $1, \ldots, n$ annehmen.

Nehmen wir also an, daß ein solcher innerer Punkt x_0 existiert, wo die Funktionen $y'(x), \ldots, y^{(n)}(x)$ stetig sind, und es eine solche Zahl $c > 0$ gibt, für die

$$(3.2) \qquad \bar{f}_{y^{(p)}y^{(p)}}(x_0) < -2c$$

gilt.

Da die Funktion $\bar{f}_{y^{(p)}y^{(p)}}(x)$ im Punkte x_0 stetig ist, gibt es eine solche Zahl δ ($x_1 < x_0 - \delta$, $x_0 + \delta < x_2$, $0 < \delta < 1$), für die

$$(3.3) \qquad \bar{f}_{y^{(p)}y^{(p)}}(x) < -c \quad \text{für} \quad x_0 - \delta \leqq x \leqq x_0 + \delta$$

ist.

Wir definieren nun eine Funktionenfolge $\eta_{k,p}(x)$ ($k = 1, 2, \ldots$) folgendermaßen:

$$(3.4) \qquad \eta_{p,k}(x) = \begin{cases} \dfrac{k}{p!}(x - x_0 + \delta)^p & \text{für} \quad x_0 - \delta \leqq x \leqq x_0 - \delta + \dfrac{\delta}{k}, \\ \dfrac{k}{p!}(-1)^p \left(x - x_0 + \delta - \dfrac{2\delta}{k}\right)^p & \text{für} \quad x_0 - \delta + \dfrac{\delta}{k} \leqq x \leqq x_0 - \delta + \dfrac{2\delta}{k}, \\ \text{sei nach } \dfrac{2\delta}{k} \text{ periodisch} & \text{für} \quad x_0 - \delta \leqq x \leqq x_0 + \delta, \\ 0 & \text{für} \quad \begin{cases} x_1 \leqq x \leqq x_0 - \delta, \\ x_0 + \delta \leqq x \leqq x_2. \end{cases} \end{cases}$$

[27] Siehe Fußnote [11].
[28] Siehe z. B. [7], S. 32.

Für alle Funktionen (3.4) gelten folgende Behauptungen:

$$(3.5)\qquad \eta_{p,k}^{(p+\alpha)}(x-0)=\eta_{p,k}^{(p+\alpha)}(x+0)=0 \quad \text{für} \quad \begin{cases} x_1<x<x_2, \\ \alpha=1,\dots,n-p, \end{cases}$$

$$(3.6)\qquad \eta_{p,k}^{(i)}(x)=0 \quad \text{für} \quad \begin{cases} x_1<x<x_0-\delta, \\ x_0+\delta<x<x_2, \\ i=0,1,\dots,p, \end{cases}$$

$$(3.7)\qquad |\eta_{p,k}^{(i)}(x-0)|<1, \quad |\eta_{p,k}^{(i)}(x+0)|<1 \quad \text{für} \quad \begin{cases} x_0-\delta<x<x_0+\delta, \\ i=0,1,\dots,p-1, \end{cases}$$

$$(3.8)\qquad \eta_{p,k}^{(p)}(x)=k \quad \text{für} \quad \begin{cases} x_0-\delta+\beta\dfrac{2\delta}{k}<x<x_0-\delta+(2\beta+1)\dfrac{\delta}{k}, \\ \beta=0,1,\dots,k-1, \end{cases}$$

$$(3.9)\qquad \eta_{p,k}^{(p)}(x)=(-1)^p k \quad \text{für} \quad \begin{cases} x_0-\delta+(2\beta+1)\dfrac{\delta}{k}<x<x_0-\delta+(\beta+1)\dfrac{2\delta}{k}, \\ \beta=0,1,\dots,k-1. \end{cases}$$

Betrachten wir nun die Funktionenschar

$$(3.10)\qquad y(x)+\varepsilon\eta_{p,k}(x) \quad \text{für} \quad k=1,2,\dots.$$

Es ist offenbar, daß die Funktionen (3.10) für hinreichend kleines $|\varepsilon|<\varepsilon_k$ $(\varepsilon_k>0)$ zur Funktionenklasse M gehören, und die Abstände $\varrho_n(y, y+\varepsilon\eta_{p,k})$ beliebig klein werden. Alle Funktionen

$$(3.11)\qquad \varphi_k(\varepsilon)=I[y(x)+\varepsilon\eta_{p,k}(x)]$$

nehmen also bei $\varepsilon=0$ ein relatives Minimum an. Das bedeutet aber, daß die Ungleichungen

$$(3.12)\qquad \varphi_k''(0)\geqq 0 \qquad (k=1,2,\dots)$$

bestehen sollen.[20]

Nach zweimaliger Differentiation bekommen wir unter Berücksichtigung von (3.5) und (3.6)

$$(3.13)\qquad \varphi_k''(0)=\int\limits_{x_0-\delta}^{x_0+\delta}\sum_{i,j=0}^{p}\bar f_{y^{(i)}y^{(j)}}\cdot\eta_{p,k}^{(i)}\eta_{p,k}^{(j)}\,dx \quad \text{für} \quad k=1,2,\dots.$$

[20] Die Funktionen $\varphi_k(\varepsilon)$ sind für $|\varepsilon|<\varepsilon_k$ auch zweimal stetig differenzierbar. Siehe z. B. [2], S. 63.

Wir erhalten für die Summe unter dem Integralzeichen unter Benützung von (3.2), (3.7), (3.8) und (3.9)

$$\begin{aligned}
&\sum_{i,j=0}^{p} \bar{f}_{y^{(i)}y^{(j)}} \cdot \eta_{p,k}^{(i)} \eta_{p,k}^{(j)} = \sum_{i,j=0}^{p-1} \bar{f}_{y^{(i)}y^{(j)}} \cdot \eta_{p,k}^{(i)} \eta_{p,k}^{(j)} + \\
(3.14)\qquad &+2\sum_{i=0}^{p-1} \bar{f}_{y^{(i)}y^{(p)}} \cdot \eta_{p,k}^{(i)} \eta_{p,k}^{(p)} + \bar{f}_{y^{(p)}y^{(p)}} \cdot \eta_{p,k}^{(p)} \eta_{p,k}^{(p)} < \\
&< \sum_{i,j=0}^{p-1} |\bar{f}_{y^{(i)}y^{(j)}}| + 2k\sum_{i=0}^{p-1} |\bar{f}_{y^{(i)}y^{(p)}}| - ck^2.
\end{aligned}$$

Da alle Funktionen $|\bar{f}_{y^{(i)}y^{(j)}}|$ im Intervall $[x_0-\delta, x_0+\delta]$ beschränkt sind, folgt aus (3.14), daß für entsprechende Konstanten

$$(3.15)\qquad \sum_{i,j=0}^{p} \bar{f}_{y^{(i)}y^{(j)}} \cdot \eta_{p,k}^{(i)} \eta_{p,k}^{(j)} < a + bk - ck^2 \quad \text{für} \quad x_0-\delta \leqq x \leqq x_0+\delta$$

besteht.[30]

(3.15) bedeutet aber, daß die Ungleichung (3.12) für beliebig großes k nicht bestehen kann.

Damit ist unsere Behauptung bewiesen.

Bemerkung. Wir betrachten jetzt solche Operationen, welche die Ungleichungen (3.1) bei allen zulässigen Funktionen so erfüllt, daß das Gleichheitszeichen überhaupt nicht vorkommt, d. h. wenn auf der Menge R

$$(3.16)\qquad f_{y^{(i)}y^{(i)}}(x, y, y', \ldots, y^{(n)}) > 0 \quad \text{für} \quad i = 1, \ldots, n$$

ist.

Nach Satz 3 kann diese Operation nur bei solchen Funktionen ein Minimum annehmen, bei denen in den Eckenpunkten — wenn solche überhaupt existieren — *wenigstens zwei Ableitungen* eine Unstetigkeit besitzen.

§ 4. Verallgemeinerung der Weierstraßschen Bedingung. Eine andere Eckenbedingung

Führen wir jetzt die folgende Funktion ein:

$$\begin{aligned}
&E(x, y, y', \ldots, y^{(n)}; p', \ldots, p^{(n)}) = f(x, y, p', \ldots, p^{(n)}) - \\
1)\qquad &-f(x, y, y', \ldots, y^{(n)}) - (p'-y') f_{y'}(x, y, y', \ldots, y^{(n)}) \\
&\text{für } (x, y) \in G;\ y^{(i)}, p^{(i)} \text{ beliebig } (i = 1, \ldots, n).
\end{aligned}$$

[30] Siehe Fußnote [11].

Die Funktion (4.1) fällt für $n=1$ mit der bekannten Weierstraßschen E-Funktion zusammen.[31]

Mit Hilfe der Funktion (4.1) können wir den folgenden Satz beweisen

SATZ 5. *Wenn die Operation* (1.12) *bei der Funktion* $y(x)\in M$ *ein* ***starkes*** *relatives Minimum annimmt, soll die Ungleichung*

$$E(x, y(x), y'(x), \ldots, y^{(n)}(x);\ p', \ldots, p^{(n)}) \geqq 0 \tag{4.2}$$

für beliebige Konstanten $p', \ldots, p^{(n)}$ *in jedem Punkte des Intervalls* $x_1 \leqq x \leqq x_2$ *bestehen.*[32]

Für $n=1$ ist Satz 5 als die Weierstraßsche notwendige Bedingung bekannt.[33]

BEWEIS. Es ist genug, den Satz wegen der Stetigkeit der Funktion (4.1) nur für solche Punkte des Intervalls $x_1 < x < x_2$ nachzuweisen, in denen alle Ableitungen $y'(x), \ldots, y^{(n)}(x)$ stetig sind.

Nehmen wir gegen die Behauptung an, daß ein solcher innerer Punkt x_0 und solche Zahlen $p_0', \ldots, p_0^{(n)}$ existieren, daß die Ungleichung

$$E(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0);\ p_0', \ldots, p_0^{(n)}) < 0 \tag{4.3}$$

gilt, und die Ableitungen $y'(x), \ldots, y^{(n)}(x)$ im Punkte x_0 stetig sind. Man kann wegen der Definition der Funktionenklasse M eine solche Zahl a $(x_1 < a < x_0)$ angeben, daß es im Intervall $a \leqq x \leqq x_0$ keinen Eckenpunkt gibt. Wir betrachten jetzt für alle hinreichend kleinen $h \geqq 0$ die Funktionen

$$y_h(x) = \begin{cases} y(x) & \text{für } \begin{cases} x_1 \leqq x \leqq a, \\ x_0 \leqq x \leqq x_2, \end{cases} \\ y(x) + k(h)\cdot(x-a) & \text{für } a \leqq x \leqq x_0 - h, \\ y(x_0) + p_0'(x-x_0) + \dfrac{p_0''}{2!}(x-x_0)^2 + & \\ \qquad + \cdots + \dfrac{p_0^{(n)}}{n!}(x-x_0)^n & \text{für } x_0 - h \leqq x \leqq x_0, \end{cases} \tag{4.4}$$

wo $k(h)$ so gewählt wurde, daß die Funktion $y_h(x)$ auch im Punkte x_0-h stetig ist; $k(h)$ soll also die Gleichung

$$\begin{aligned} & y(x_0-h) + k(h)\cdot(x_0-h-a) = \\ & = y(x_0) - p_0'h + \frac{1}{2!}p_0''h^2 - \cdots + (-1)^n \frac{1}{n!}p_0^{(n)}h^n \end{aligned} \tag{4.5}$$

[31] Siehe z. B. [2], S. 110.

[32] Siehe Fußnote [11].

[33] Siehe z. B. [3], S. 131.

erfüllen. Man kann aus (4. 5)

$$k(0)=0, \quad k'(0)=-\frac{p_0'-y'(x_0)}{x_0-a} \tag{4.6}$$

leicht gewinnen.

Es ist klar, daß die Funktionen $y_h(x)$ für hinreichend kleines $h \geqq 0$ zur Funktionenklasse M gehören, und die Abstände $\varrho(y, y_h)$ beliebig klein werden.

Untersuchen wir nachher die Funktion

$$\begin{aligned} \Delta(h) &= I[y_h(x)]-I[y(x)]= \\ &= \int_a^{x_0-h} f(x, y(x)+k(h)\cdot(x-a), y'(x)+k(h), y''(x), \ldots, y^{(n)}(x))dx+ \\ &+\int_{x_0-h}^{x_0} f\Big(x, y(x_0)+\sum_{l=1}^{n}\frac{1}{l!}p_0^{(l)}(x-x_0)^l, \sum_{l=1}^{n}\frac{1}{(l-1)!}p_0^{(l)}(x-x_0)^{l-1}, \ldots, p_0^{(n)}\Big)dx- \\ &-\int_a^{x_0} f(x, y(x), y'(x), \ldots, y^{(n)}(x))dx. \end{aligned} \tag{4.7}$$

Es ist leicht einzusehen, daß die Funktion $\Delta(h)$ für hinreichend kleines $h \geqq 0$ stetig differenzierbar ist.[34]

Nach einer einfachen Rechnung bekommen wir aus (4. 7) unter Benützung von (4. 6) und von der Gleichung (2. 50) folgendes:

$$\begin{aligned} \Delta'(+0) &= -f(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0))+ \\ &+\int_a^{x_0}\{f_y(x, y(x), y'(x), \ldots, y^{(n)}(x))\cdot k'(0)\cdot(x-a)+ \\ &+f_{y'}(x, y(x), y'(x), \ldots, y^{(n)}(x))\cdot k'(0)\}\,dx+f(x_0, y(x_0), p_0', \ldots, p_0^{(n)})= \\ &=f(x_0, y(x_0), p_0', \ldots, p_0^{(n)})-f(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0))+ \\ &+k'(0)\int_a^{x_0}\Big\{f_{y'}(x, y(x), y'(x), \ldots, y^{(n)}(x))+ \\ &+(x-a)\frac{d}{dx}f_{y'}(x, y(x), y'(x), \ldots, y^{(n)}(x))\Big\}\,dx= \\ &=f(x_0, y(x_0), p_0', \ldots, p_0^{(n)})-f(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0))+ \\ &+k'(0)\int_a^{x_0}\frac{d}{dx}\{(x-a)\cdot f_{y'}(x, y(x), y'(x), \ldots, y^{(n)}(x))\}\,dx= \end{aligned} \tag{4.8}$$

[34] Siehe z. B. [2], S. 63. Für $h=0$ soll nur die rechtsseitige Derivierte existieren.

$$=f(x_0, y(x_0), p_0', \ldots, p_0^{(n)})-f(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0)) -$$
$$-\frac{p_0'-y'(x_0)}{x_0-a}\,[(x-a)f_{y'}(x, y(x), y'(x), \ldots, y^{(n)}(x))]_a^{x_0}=$$
$$=f(x_0, y(x_0); p_0', \ldots, p_0^{(n)})-f(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0))-$$
$$-(p_0'-y'(x_0))f_{y'}(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0))=$$
$$=E(x_0, y(x_0), y'(x_0), \ldots, y^{(n)}(x_0); p_0', \ldots, p_0^{(n)}).$$

Wir erhielten also unter Berücksichtigung von (4.3)

(4.9) $$\Delta'(+0)<0.$$

Da $\Delta(0)=0$, folgt aus (4.9), daß für alle hinreichend kleinen $h>0$

(4.10) $$\Delta(h)=I[y_h(x)]-I[y(x)]<0$$

gilt.

(4.10) widerspricht aber offensichtlich der Tatsache, daß die Operation (1.12) bei der Funktion $y(x)$ ein starkes relatives Minimum annimmt.

Der Satz ist also bewiesen.

Folgen und Bemerkungen.

1. Eine Folge der Bedingung (4.2) ist der

SATZ 6. *Wenn die Operation* (1.12) *bei der Funktion* $y(x)\in M$ *ein starkes relatives Minimum annimmt, soll die Funktion* $y(x)$ *die Eckenbedingung*

(4.11) $$\begin{aligned}&f(x, y(x), y'(x-0), \ldots, y^{(n)}(x-0))-\\ &-y'(x-0)f_{y'}(x, y(x), y'(x-0), \ldots, y^{(n)}(x-0))=\\ &=f(x, y(x), y'(x+0), \ldots, y^{(n)}(x+0))-\\ &-y'(x+0)f_{y'}(x, y(x), y'(x+0), \ldots, y^{(n)}(x+0))\end{aligned}$$

in jedem Punkte des Intervalls $x_1<x<x_2$ *erfüllen.*

Für $n=1$ fallen (4.11) und (2.2) mit den Weierstraß—Erdmannschen Eckenbedingungen zusammen.[35]

BEWEIS. Wir sollen den Beweis offenbar nur für die Eckenpunkte der Funktion $y(x)$ durchführen. Sei also x_0 ein beliebiger Eckenpunkt. Wir nehmen zuerst die linksseitigen Grenzwerte der Ableitungen der Funktion $y(x)$ im Punkte x_0, und wählen für die Zahlen $p^{(i)}$ $(i=1, \ldots, n)$ die rechtsseitigen Grenzwerte $y^{(i)}(x_0+0)$ $(i=1, \ldots, n)$. Dann lautet die Bedingung (4.2)

(4.12) $$E(x_0, y(x_0), y'(x_0-0), \ldots, y^{(n)}(x_0-0);\; y'(x_0+0), \ldots, y^{(n)}(x_0+0))\geqq 0.$$

[35] Siehe z. B. [2], S. 367.

Vertauschen wir jetzt die Rolle der Zahlen $y^{(i)}(x_0-0)$ und $y^{(i)}(x_0+0)$ $(i=1,\dots,n)$. Ebenso gewinnen wir aus (4. 2)

$$E(x_0, y(x_0), y'(x_0+0),\dots, y^{(n)}(x_0+0);\, y'(x_0-0),\dots, y^{(n)}(x_0-0)) \geqq 0. \tag{4.13}$$

Es folgt aber wegen der Eckenbedingung (2. 2) aus der Gestalt der Funktion (4. 1)

$$\begin{aligned} &E(x_0, y(x_0), y'(x_0-0),\dots y^{(n)}(x_0-0);\, y'(x_0+0),\dots, y^{(n)}(x_0+0)) = \\ &= -E(x_0, y(x_0), y'(x_0+0),\dots, y^{(n)}(x_0+0);\, y'(x_0-0),\dots, y^{(n)}(x_0-0)), \end{aligned} \tag{4.14}$$

also aus (4. 12) und (4. 13) ergibt sich

$$E(x_0, y(x_0), y'(x_0-0),\dots, y^{(n)}(x_0-0);\, y'(x_0+0),\dots, y^{(n)}(x_0+0)) = 0. \tag{4.15}$$

Wenn wir (4. 15) ausführlich ausschreiben und die Eckenbedingung (2. 2) wieder benützen, so bekommen wir eben unsere Behauptung.

2. Satz 5 — wie dies sich aus der obigen Beweisführung ergibt — verändert sich im Falle eines schwachen relativen Minimums folgendermaßen:

ZUSATZ 1. *Wenn die Operation bei der Funktion $y(x)\in M$ ein schwaches relatives Minimum annimmt, gibt es eine Zahl $\varepsilon>0$, so daß die Ungleichung*

$$E(x, y(x), y'(x),\dots, y^{(n)}(x);\, p',\dots, p^{(n)}) \geqq 0 \tag{4.16}$$

$$\text{für}\quad |p^{(i)}-y^{(i)}(x)|<\varepsilon \qquad (i=1,\dots,n)$$

in jedem Stetigkeitspunkte[36] *des Intervalls $x_1<x<x_2$ besteht.*

3. Nehmen wir an, daß die Funktion $y(x)\in M$ der Bedingung (4. 16) genügt. Ferner sei x_0 ein beliebiger Stetigkeitspunkt der Funktion $y(x)$ im Intervall $x_1<x<x_2$. Aus der Bedingung (4. 16) folgt, daß die Funktion $E(x_0, y(x_0), y'(x_0),\dots, y^{(n)}(x_0);\, p',\dots, p^{(n)})$ bei $p^{(i)}=y^{(i)}(x_0)$ $(i=1,\dots,n)$ ein relatives Minimum annimmt. Es sollen also die Relationen

$$f_{y^{(i)}}(x_0, y(x_0), y'(x_0),\dots, y^{(n)}(x_0)) = 0 \quad \text{für}\quad i=2,\dots,n, \tag{4.17}$$

$$f_{y^{(i)}y^{(i)}}(x_0, y(x_0), y'(x_0),\dots, y^{(n)}(x_0)) \geqq 0 \quad \text{für}\quad i=1,\dots,n \tag{4.18}$$

und

$$\sum_{i,j=1}^{n} f_{y^{(i)}y^{(j)}}(x_0, y(x_0), y'(x_0),\dots, y^{(n)}(x_0))\sigma_i\sigma_j \geqq 0 \tag{4.19}$$

$$\text{für beliebige}\quad \sigma_1,\dots,\sigma_n$$

bestehen.

36 Darunter verstehen wir solche Punkte, in denen die Ableitungen $y'(x),\dots,y^{(n)}(x)$ stetig sind.

Damit haben wir folgendes bewiesen:

ZUSATZ 2. *Wenn irgendwelche Funktion $y(x) \in M$ die Bedingung* (4.16) *erfüllt, so genügt sie auch den Gleichungen* (2.1) *und den Bedingungen* (3.1).[37]

ZUSATZ 3. *Wenn die Operation* (1.12) *bei der Funktion $y(x) \in M$ ein schwaches relatives Minimum annimmt, soll die Funktion $y(x)$ für beliebige Zahlen $\sigma_1, \ldots, \sigma_n$ die Ungleichung*

$$\sum_{i,j=1}^{n} f_{y^{(i)}y^{(j)}}(x, y(x), y'(x), \ldots, y^{(n)}(x))\sigma_i\sigma_j \geqq 0 \tag{4.20}$$

in jedem Stetigkeitspunkte des Intervalls $x_1 < x < x_2$ erfüllen.

4. *Alles zusammengefaßt, können folgende festgestellt werden: Alle von uns abgeleiteten notwendigen Bedingungen vom neuen Typ* ((2.1) *und* (3.1) *für $n \geqq 2$, $i = 2, \ldots, n$) ergeben sich aus der Bedingung* (4.2).[38] *Die Bedingung* (4.2) *ist aber — hinsichtlich der Form der in ihr vorhandenen E-Funktion — der im Falle $n=1$ wohlbekannten Weierstraßschen notwendigen Bedingung völlig ähnlich. Die bisher für $n=1$ bekannten notwendigen Bedingungen*[39] *— in ihrer unveränderten Form — sind also die charakteristischen notwendigen Bedingungen nicht für den Fall $n=1$, sondern im allgemeinen für die diskontinuierlichen Probleme.*

(Eingegangen am 1. Juli 1959.)

Literaturverzeichnis

[1] A. KNESER, *Lehrbuch der Variationsrechnung* (1900).
[2] O. BOLZA, *Vorlesungen über Variationsrechnung* (Leipzig, 1949).
[3] G. A. BLISS, *Calculus of variations* (Chicago, 1925).
[4] G. A. BLISS, *Lectures on the calculus of variations* (Chicago, 1945).
[5] C. CARATHÉODORY, *Über die diskontinuierlichen Lösungen in der Variationsrechnung*, Dissertation (Göttingen, 1904), Gesammelte math. Schriften, Band I (München, 1954).
[6] H. H. PIXLEY, A problem in the calculus of variations suggested by a problem in economics, *Contributions to calculus of variations*, (1931—1932), S. 131—189.
[7] M. PICONE, Sulle condizioni necessarie per un estremo, nel calcolo delle variazioni, *Publ. dell'Ist. per le Appl. del Calcolo*, N. 410 (Roma, 1954).
[8] E. E. LEVI, Sui criterie sufficienti per il massimo e per il minimo nel calcolo delle variazioni, *Annali di Mat.*, (3) **21**, S. 173—218.
[9] Н. М. Гюнтер, Курс вариационного исчисления (Москва—Ленинград, 1941).
[10] Н. И. Ахиеэер, Лекции по вариационному исчислению (Москва, 1955).

[37] Natürlich ist die Bedingung (4.16) in der Bedingung (4.2) enthalten.

[38] Im Falle eines schwachen relativen Minimums aus der Bedingung (4.16). Man muß natürlich beachten, daß die Gleichung (2.50) bei der Herleitung der Bedingung (4.2) ausgenützt worden war. Wir haben aber die Gleichung (2.50) mit Hilfe der Gleichungen (2.1) hergeleitet.

[39] D. h. die Euler—Lagrangesche Gleichung, die Legendresche Bedingung, die Weierstraßsche Bedingung und die Weierstraß—Erdmannschen Eckenbedingungen.

ЗАМЕЧАНИЯ ОБ ИНТЕРПОЛИРОВАНИИ

О. КИШ (Будапешт)

(Представлено П. Тураном)

§ 1

Настоящая заметка примыкает к ряду работ, посвященных теории (0, 2)-интерполяционных многочленов. Так инициатор изучения этих многочленов П. Туран назвал многочлены $R_n(x)$ не выше $2n-1$-ой степени, удовлетворяющие в некоторых узлах $x_1, x_2, \ldots, x_n$ условиям

$$R_n(x_k)=\alpha_k, \quad R_n''(x_k)=\beta_k \qquad (k=1,2,\ldots,n),$$

где α_k и β_k некоторые вещественные числа.

Приведем некоторые относящиеся сюда результаты.

Если узлы интерполирования расположены симметрично относительно нуля и число их нечетно, то (0, 2)-интерполяционный многочлен, вообще говоря, не существует. Если узлы суть корни ультрасферического многочлена $P_n^{(\lambda)}(x)$, $n \geqq 4$ четно, $\lambda \geqq -\frac{1}{2}$ и $\lambda+\frac{1}{2}$ не есть четное натуральное число, то всегда существует единственный (0, 2)-интерполяционный многочлен.[1]

Для случая, когда узлы суть корни многочлена $\Pi_n(x)=(1-x^2)P_{n-1}'(x)$, где $P_n(x)$ — многочлен Лежандра, и число их четно, получен явный вид (0, 2)-интерполяционных многочленов.[2]

Если для определенной на отрезке $[-1, +1]$ функции $f(x)$ выполняется условие

$$f(x+h)-2f(x)+f(x-h)=o(h),$$

узлы суть корни многочлена $\Pi_n(x)$,

$$\alpha_k=f(x_k) \qquad (k=1,2,\ldots,n)$$

и β_k не слишком велики, то последовательность (0, 2)-интерполяционных многочленов равномерно сходится к $f(x)$ на отрезке $[-1, +1]$.[3]

[1] J. Surányi and P. Turán, Notes on interpolation. I, *Acta Math. Acad. Sci. Hung.*, **6** (1955), стр. 67—79.

[2] J. Balázs and P. Turán, Notes on interpolation. II, *Acta Math. Acad. Sci. Hung.*, **8** (1957), стр. 201—215.

[3] G. Freud, Bemerkungen über die Konvergenz eines Interpolationsverfahrens von P. Turán, *Acta Math. Acad. Sci. Hung.*, **9** (1958), стр. 337—341.

Как показывает следующая теорема, этот результат является в некотором смысле наилучшим. Пусть $0<\varepsilon<1$, $\beta_k=0$. Существует функция $f(x)$, удовлетворяющая условию Липшица с показателем $1-\varepsilon$, для которой последовательность $(0, 2)$-интерполяционных многочленов неограниченна в точке 0.[4]

Из сказанного следует, что условие сходимости $(0, 2)$-интерполяционных многочленов значительно более жесткое чем условие сходимости интерполяционных многочленов Лагранжа при тех же узлах.

Результаты, аналогичные приведенным выше, имеют место и для $(0, 1, 3)$-интерполяционных многочленов, когда в узлах заданы значения многочлена не выше $3n-1$-ой степени и его первой и третьей производной.[5, 6]

Ниже исследуются $(0, 2)$-интерполяционные многочлены в том случае, когда узлы интерполирования суть

$$z_k=\exp i\frac{2\pi}{n}k \qquad (k=1, 2, \ldots, n).$$

Оказывается, что при $n\geqq 2$ $(0, 2)$-интерполяционный многочлен всегда существует и единственен. Таким образом, в этом случае, в отличии от упомянутого выше, безразлично является ли число узлов четным или нечетным.

Дается явный вид этих многочленов, после чего изучается их сходимость. Оказывается, что если функция $f(z)$ регулярна при $|z|<1$ и непрерывна при $|z|\leqq 1$, модуль непрерывности функции $f(\exp ix)$ удовлетворяет условию Дини—Липшица,

$$\alpha_k=f(z_k) \qquad (k=1, 2, \ldots, n)$$

и β_k не слишком велики, то $(0, 2)$-интерполяционные многочлены равномерно сходятся к $f(z)$ в круге $|z|\leqq 1$.

Как известно, интерполяционный процесс Лагранжа сходится при этих же условиях, что, как мы видели, не имеет место в случае, когда узлы суть корни многочлена $\Pi_n(x)$.

Далее дается явный вид $(0, 1, \ldots, r-2, r)$-интерполяционных многочленов. Случай $r=2$ рассматривался отдельно для того, чтобы на этом простейшем примере можно было бы лучше уяснить сущность дела. Наконец, изучается сходимость $(0, 1, 3)$-интерполяционных многочленов.

[4] J. Balázs and P. Turán, Notes on interpolation. III, *Acta Math. Acad. Sci. Hung.*, 9 (1958), стр. 195—217.

[5] R. B. Saxena and A. Sharma, On some interpolatory properties of Legendre polynomials, *Acta Math. Acad. Sci. Hung.*, 9 (1958), стр. 345—358.

[6] R. B. Saxena and A. Sharma, Convergence of interpolatory polynomials, *Acta Math. Acad. Sci. Hung.*, 10 (1959), стр. 157—175.

§ 2

В этом §-е приводится явный вид (0, 2)-интерполяционных многочленов и доказывается их единственность.

Теорема 1. Если

$$z_k = \exp i \frac{2\pi}{n} k \qquad (k = 1, 2, \ldots, n;\ n \geqq 2), \tag{1}$$

$$l_k(z) = \frac{z_k}{n} \frac{z^n - 1}{z - z_k}, \tag{2}$$

$$A_k(z) = l_k(z) - \frac{1}{2n}\left(z^{\frac{n+1}{2}} - z^{-\frac{n-1}{2}}\right)\int_0^z t^{\frac{n+1}{2}} l_k''(t)\,dt, \tag{3}$$

$$B_k(z) = \frac{z_k^2}{2n}\left(z^{\frac{n+1}{2}} - z^{-\frac{n-1}{2}}\right)\int_0^z t^{\frac{n-3}{2}} l_k(t)\,dt, \tag{4}$$

α_k и β_k суть любые комплексные числа, то функция

$$R_n(z) = \sum_{k=1}^{n} \alpha_k A_k(z) + \sum_{k=1}^{n} \beta_k B_k(z) \tag{5}$$

есть единственный многочлен не выше $2n-1$-ой степени, удовлетворяющий условиям

$$R_n(z_k) = \alpha_k \qquad (k = 1, 2, \ldots, n), \tag{6}$$

$$R_n''(z_k) = \beta_k \qquad (k = 1, 2, \ldots, n). \tag{7}$$

Доказательство. Докажем сначала, что функции $A_k(z)$ являются многочленами $2n-1$-ой степени. Так как в силу (2) и (1)

$$l_k(z) = \frac{1}{n} \sum_{\nu=0}^{n-1} z_k^{n-\nu} z^\nu, \tag{8}$$

то

$$t^{\frac{n+1}{2}} l_k''(t) = \frac{1}{n} \sum_{\nu=2}^{n-1} \nu(\nu-1)\, z_k^{n-\nu}\, t^{\frac{n-3}{2}+\nu},$$

$$\int_0^z t^{\frac{n+1}{2}} l_k''(t)\,dt = \frac{2}{n} \sum_{\nu=2}^{n-1} \frac{\nu(\nu-1)}{n-1+2\nu} z_k^{n-\nu} z^{\frac{n-1}{2}+\nu},$$

$$\left(z^{\frac{n+1}{2}} - z^{-\frac{n-1}{2}}\right)\int_0^z t^{\frac{n+1}{2}} l_k''(t)\,dt = \frac{2}{n} \sum_{\nu=2}^{n-1} \frac{\nu(\nu-1)}{n-1+2\nu} z_k^{n-\nu}(z^{n+\nu} - z^\nu). \tag{9}$$

Из (3), (8) и (9) следует, что $A_k(z)$ действительно многочлены $2n-1$-ой степени.

Аналогичным образом доказывается, что и $B_k(z)$ суть многочлены $2n-1$-ой степени.

Из сказанного и из (5) следует, что $R_n(z)$ действительно есть многочлен не выше $2n-1$-ой степени.

Вычислим значения функции

$$\omega(z)=z^{\frac{n+1}{2}}-z^{-\frac{n-1}{2}} \tag{10}$$

и ее первой и второй производной в узлах, так как они понадобятся нам в дальнейшем.

Очевидно

$$\omega(z)=z^{-\frac{n-1}{2}}(z^n-1).$$

Отсюда и из следующего из (1) равенства

$$z_k^n=1 \tag{11}$$

получаем

$$\omega(z_k)=0 \qquad (k=1,2,\ldots,n). \tag{12}$$

Дифференцируя (10), имеем

$$\omega'(z)=\frac{n+1}{2}z^{\frac{n-1}{2}}+\frac{n-1}{2}z^{-\frac{n+1}{2}}=z^{-\frac{n+1}{2}}\left(\frac{n+1}{2}z^n+\frac{n-1}{2}\right). \tag{13}$$

Отсюда и из (11) следует

$$\omega'(z_k)=z_k^{-\frac{n+1}{2}}\left(\frac{n+1}{2}z_k^n+\frac{n-1}{2}\right)=z_k^{-\frac{n+1}{2}}\left(\frac{n+1}{2}+\frac{n-1}{2}\right)=nz_k^{-\frac{n+1}{2}}. \tag{14}$$

Дифференцируя (13), получаем

$$\omega''(z)=\frac{n+1}{2}\,\frac{n-1}{2}z^{\frac{n-3}{2}}-\frac{n-1}{2}\,\frac{n+1}{2}z^{-\frac{n+3}{2}}=\frac{n+1}{2}\,\frac{n-1}{2}z^{-\frac{n+3}{2}}(z^n-1).$$

Отсюда и из (11) следует равенство

$$\omega''(z_k)=\frac{n+1}{2}\,\frac{n-1}{2}z_k^{-\frac{n+3}{2}}(z_k^n-1)=0. \tag{15}$$

Переходим к доказательству соотношений (6) и (7). Для фундаментальных многочленов Лагранжева интерполирования $l_k(z)$ имеют место равенства

$$l_k(z_j)=\begin{cases}1, & \text{если } j=k\\ 0, & \text{если } j\neq k\end{cases} \qquad (j,k=1,2,\ldots,n).$$

Отсюда, из (3), (4), (10) и (12) следуют соотношения

(16) $$A_k(z_j) = \begin{cases} 1, & \text{если } j = k \\ 0, & \text{если } j \neq k \end{cases} \qquad (j, k = 1, 2, \ldots, n),$$

(17) $$B_k(z_j) = 0 \qquad (j, k = 1, 2, \ldots, n).$$

Вычислим вторую производную $A_k(z)$, дифференцируя (3) и применяя при этом правило Лейбница:

$$A_k''(z) = l_k''(z) - \frac{1}{2n}\left\{\omega''(z)\int_0^z t^{\frac{n+1}{2}} l_k''(t)\,dt + 2\omega'(z)\, z^{\frac{n+1}{2}} l_k''(z) + \omega(z)\,[z^{\frac{n+1}{2}} l_k''(z)]'\right\}.$$

Отсюда, из (12), (14) и (15) следует

(18) $$A_k''(z_j) = l_k''(z_j) - \frac{1}{2n}\, 2n z_j^{-\frac{n+1}{2}} z_j^{\frac{n+1}{2}} l_k''(z_j) = 0 \qquad (j, k = 1, 2, \ldots, n).$$

Аналогичным образом можно показать, что

(19) $$B_k''(z_j) = \begin{cases} 1, & \text{если } j = k \\ 0, & \text{если } j \neq k \end{cases} \qquad (j, k = 1, 2, \ldots, n).$$

Из (5) и (16)—(19) следует, что условия (6) и (7) действительно выполняются.

Остается доказать, что полученный многочлен есть единственный многочлен, удовлетворяющий этим условиям.

Пусть

$$R_n(z) = \sum_{\nu=0}^{2n-1} c_\nu z^\nu .$$

Условия (6) и (7) можно записать в виде

$$\sum_{\nu=0}^{2n-1} z_k^\nu c_\nu = \alpha_k \qquad (k = 1, 2, \ldots, n),$$

$$\sum_{\nu=2}^{2n-1} \nu(\nu-1)\, z_k^{\nu-2} c_\nu = \beta_k \qquad (k = 1, 2, \ldots, n).$$

Мы видели, что эта система уравнений относительно c_ν всегда имеет решение, так как всегда существует многочлен не выше $2n-1$-ой степени, удовлетворяющий условиям (6) и (7). Отсюда следует, что ее определитель не равен нулю. Но тогда она может иметь лишь единственное решение, т. е. наш (0, 2)-интерполяционный многочлен единственен, что и требовалось доказать.

§ 3

В этом §-е доказывается одна лемма об интеграле Джексона и, используя ее, еще одно вспомогательное предложение, которое понадобится в дальнейшем. Возможно, они не новы, но я не могу сослаться на работы, где они были бы опубликованы.

Лемма 1. Пусть $g(x)$ есть комплексная, 2π-периодическая, непрерывная функция вещественного аргумента, $\omega(\delta)$ ее модуль непрерывности. Для интеграла Джексона

$$U_n(x) = \frac{3}{2\pi n(2n^2+1)} \int_{-\pi}^{+\pi} g(t) \left[\frac{\sin n \frac{t-x}{2}}{\sin \frac{t-x}{2}} \right]^4 dt \tag{20}$$

имеет место неравенство

$$|U_n'(x)| \leqq 20 n \omega\left(\frac{1}{n}\right). \tag{21}$$

Доказательство. Чтобы упростить выкладки, введем обозначение

$$D_n(x) = \frac{3}{2\pi n(2n^2+1)} \left(\frac{\sin nx}{\sin x} \right)^4. \tag{22}$$

Дифференцируя (20) и полагая $t = x + 2s$, имеем

$$U_n'(x) = -\frac{1}{2} \int_{-\pi}^{+\pi} g(t) D_n'\left(\frac{t-x}{2}\right) dt = -\int_{-\frac{\pi}{2}-\frac{x}{2}}^{+\frac{\pi}{2}-\frac{x}{2}} g(x+2s) D_n'(s)\, ds.$$

Так как подинтегральная функция π-периодична, то можно изменить границы интегрирования:

$$U_n'(x) = -\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} g(x+2s) D_n'(s)\, ds.$$

Разобьем этот интеграл на два, положим в первом $s = -t$, а во втором $s = t$, а затем, принимая во внимание нечетность функции $D'(t)$, сложим их:

$$U_n'(x) = -\int_{-\frac{\pi}{2}}^{0} g(x+2s) D_n'(s)\, ds - \int_{0}^{\frac{\pi}{2}} g(x+2s) D_n'(s)\, ds =$$

$$=\int_{\frac{\pi}{2}}^{0} g(x-2t)\,D_n'(-t)\,dt-\int_{0}^{\frac{\pi}{2}} g(x+2t)\,D_n'(t)\,dt=\int_{0}^{\frac{\pi}{2}} g(x-2t)\,D_n'(t)\,dt-$$

$$-\int_{0}^{\frac{\pi}{2}} g(x+2t)\,D_n'(t)\,dt=\int_{0}^{\frac{\pi}{2}}[g(x-2t)-g(x+2t)]\,D_n'(t)\,dt.$$

Наконец, разобьем и этот интеграл на два:

$$U_n'(x)=\int_{0}^{\frac{\pi}{2n}}[g(x-2t)-g(x+2t)]\,D_n'(t)\,dt+ \tag{23}$$

$$+\int_{\frac{\pi}{2n}}^{\frac{\pi}{2}}[g(x-2t)-g(x+2t)]\,D_n'(t)\,dt.$$

Оценим стоящие справа интегралы.

Начнем с первого. Так как здесь

$$0\leqq t\leqq\frac{\pi}{2n},$$

то

$$|g(x-2t)-g(x+2t)|\leqq\omega(4t)\leqq\omega\left(\frac{2\pi}{n}\right)\leqq\omega\left(\frac{7}{n}\right)\leqq 7\,\omega\left(\frac{1}{n}\right). \tag{24}$$

Ввиду того, что $|\sin nt|\leqq n|\sin t|$,

$$\left(\frac{\sin nt}{\sin t}\right)^4\leqq n^4.$$

Слева здесь стоит тригонометрический многочлен $4(n-1)$-ого порядка, поэтому в силу неравенства Бернштейна

$$\left|\left[\left(\frac{\sin nt}{\sin t}\right)^4\right]'\right|\leqq 4(n-1)\,n^4<4n^5.$$

Отсюда и из (22) следует, что

$$|D_n'(t)|\leqq\frac{3}{2\pi n(2n^2+1)}\,4n^5<\frac{3n^2}{\pi}. \tag{25}$$

Наконец, из (24) и (25) получаем

$$\left|\int_{0}^{\frac{\pi}{2n}}[g(x-2t)-g(x+2t)]\,D_n'(t)\,dt\right|\leqq 7\,\omega\left(\frac{1}{n}\right)\frac{3n^2}{\pi}\,\frac{\pi}{2n}=\frac{21}{2}\,n\,\omega\left(\frac{1}{n}\right). \tag{26}$$

Переходим к оценке второго интеграла. Теперь

$$\frac{\pi}{2n} \leqq t \leqq \frac{\pi}{2}$$

и, следовательно,

(27) $|g(x-2t)-g(x+2t)| \leqq \omega(4t) \leqq (4tn+1)\,\omega\left(\frac{1}{n}\right) \leqq \left(1+\frac{1}{2\pi}\right) 4nt\,\omega\left(\frac{1}{n}\right).$

Так как

$$\left[\left(\frac{\sin nt}{\sin t}\right)^4\right]' = 4\left(\frac{\sin nt}{\sin t}\right)^3 \frac{n\cos nt \sin t - \sin nt \cos t}{\sin^2 t}$$

и

$$|\sin nt| \leqq 1, \quad |\cos nt| \leqq 1, \quad |\cos t| \leqq 1, \quad \sin t \geqq \frac{2}{\pi} t,$$

то

$$\left|\left[\left(\frac{\sin nt}{\sin t}\right)^4\right]'\right| \leqq 4\left(\frac{n}{\sin^4 t} + \frac{1}{\sin^5 t}\right) \leqq \frac{\pi^4}{4}\left(\frac{n}{t^4} + \frac{\pi}{2}\,\frac{1}{t^5}\right).$$

Отсюда и из (22) следует

(28) $$|D_n'(t)| \leqq \frac{3}{2\pi n(2n^2+1)}\,\frac{\pi^4}{4}\left(\frac{n}{t^4} + \frac{\pi}{2}\,\frac{1}{t^5}\right) < \frac{3\pi^3}{16n^2}\left(\frac{1}{t^4} + \frac{\pi}{2n}\,\frac{1}{t^5}\right).$$

В силу (27) и (28)

$$\left|\int_{\frac{\pi}{2n}}^{\frac{\pi}{2}} [g(x-2t)-g(x+2t)]\, D_n'(t)\, dt\right| \leqq$$

$$\leqq \left(1+\frac{1}{2\pi}\right) 4n\,\omega\left(\frac{1}{n}\right) \frac{3\pi^3}{16n^2} \int_{\frac{\pi}{2n}}^{\frac{\pi}{2}} \left(\frac{1}{t^3} + \frac{\pi}{2n}\,\frac{1}{t^4}\right) dt \leqq$$

$$\leqq (2\pi+1)\,\frac{3\pi^2}{8}\,\frac{1}{n}\,\omega\left(\frac{1}{n}\right) \int_{\frac{\pi}{2n}}^{\infty} \left(\frac{1}{t^3} + \frac{\pi}{2n}\,\frac{1}{t^4}\right) dt.$$

Так как

$$\int_{\frac{\pi}{2n}}^{\infty} \left(\frac{1}{t^3} + \frac{\pi}{2n}\,\frac{1}{t^4}\right) dt = \frac{1}{2}\left(\frac{2n}{\pi}\right)^2 + \frac{1}{3}\,\frac{\pi}{2n}\left(\frac{2n}{\pi}\right)^3 = \frac{10}{3}\,\frac{n^2}{\pi^2},$$

то окончательно получаем

$$(29)\quad \left|\int\limits_{\frac{\pi}{2n}}^{\frac{\pi}{2}} [g(x-2t)-g(x+2t)]\, D_n'(t)\, dt\right| \leqq (2\pi+1)\frac{3\pi^2}{8}\frac{1}{n}\,\omega\left(\frac{1}{n}\right)\frac{10}{3}\frac{n^2}{\pi^2} =$$

$$= \frac{5(2\pi+1)}{4}\, n\omega\left(\frac{1}{n}\right) \leqq \frac{19}{2}\, n\omega\left(\frac{1}{n}\right).$$

Из (23), (26) и (29) получаем неравенство

$$|U_n'(x)| \leqq \frac{21}{2}\, n\omega\left(\frac{1}{n}\right) + \frac{19}{2}\, n\omega\left(\frac{1}{n}\right) = 20\, n\omega\left(\frac{1}{n}\right),$$

что и требовалось доказать.

Лемма 2. Пусть функция $f(z)$ регулярна при $|z|<1$ и непрерывна при $|z|\leqq 1$. Обозначим через $\omega(\delta)$ модуль непрерывности функции $f(\exp ix)$. Тогда функция

$$(30)\qquad F_n(z) = \frac{3}{2\pi n(2n^2+1)i}\, z^{2-2n} \int\limits_{|t|=1} f(t)\, t^{1-2n}\left(\frac{t^n-z^n}{t-z}\right)^4 dt$$

есть многочлен не выше $2n-2$-ой степени, удовлетворяющий в круге $|z|\leqq 1$ условиям

$$(31)\qquad |f(z)-F_n(z)| \leqq 6\omega\left(\frac{1}{n}\right),$$

$$(32)\qquad |F_n^{(m)}(z)| \leqq 10(2n)^m\, \omega\left(\frac{1}{n}\right) \qquad (m=1, 2, 3, \ldots).$$

Доказательство. Сначала докажем, что $F_n(z)$ есть многочлен не выше $2n-2$-ой степени. Так как

$$\frac{t^n-z^n}{t-z} = \sum_{k=0}^{n-1} t^k z^{n-k-1},$$

то с некоторыми коэффициентами c_k имеет место равенство

$$\left(\frac{t^n-z^n}{t-z}\right)^4 = \sum_{k=0}^{4n-4} c_k t^k z^{4n-k-4}.$$

Принимая во внимание, что в силу условий теоремы

$$\int\limits_{|t|=1} f(t)\, t^k\, dt = 0 \qquad (k=0, 1, 2 \ldots),$$

отсюда следует, что

$$\int\limits_{|t|=1} f(t)t^{1-2n}\left(\frac{t^n-z^n}{t-z}\right)^4 dt=\int\limits_{|t|=1} f(t)\,t^{1-2n}\sum_{k=0}^{4n-4} c_k t^k z^{4n-k-4}\,dt=$$

$$=\sum_{k=0}^{4n-4} c_k z^{4n-k-4}\int\limits_{|t|=1} f(t)\,t^{k-2n+1}\,dt=\sum_{k=0}^{2n-2} c_k z^{4n-k-4}\int\limits_{|t|=1} f(t)\,t^{k-2n+1}\,dt.$$

А из этого равенства и из (30) получаем

$$F_n(z)=\frac{3}{2\pi n(2n^2+1)i}\sum_{k=0}^{2n-2} c_k z^{2n-k-2}\int\limits_{|t|=1} f(t)\,t^{k-2n+1}\,dt,$$

следовательно $F_n(z)$ действительно есть многочлен не выше $2n-2$-ой степени.

Покажем теперь, что $F_n(\exp ix)$ есть интеграл Джексона функции $f(\exp ix)$. Для этого положим

$$z=\exp ix,\qquad t=\exp iy.$$

Используя равенство Эйлера, имеем

$$z^{2-2n}\,t^{2-2n}\left(\frac{t^n-z^n}{t-z}\right)^4=\left(\frac{t^{\frac{n}{2}}z^{-\frac{n}{2}}-t^{-\frac{n}{2}}z^{\frac{n}{2}}}{t^{\frac{1}{2}}z^{-\frac{1}{2}}-t^{-\frac{1}{2}}z^{\frac{1}{2}}}\right)^4=$$

$$=\left(\frac{\exp in\dfrac{y-x}{2}-\exp i\dfrac{x-y}{2}}{\exp i\dfrac{y-x}{2}-\exp i\dfrac{x-y}{2}}\right)^4=\left(\frac{\sin n\dfrac{y-x}{2}}{\sin\dfrac{y-x}{2}}\right)^4.$$

Кроме того

$$\frac{dt}{it}=\frac{1}{it}\frac{dt}{dy}dy=\frac{1}{it}it\,dy=dy.$$

Из (30) и трех последних равенств следует

$$(33)\qquad F_n(\exp ix)=\frac{3}{2\pi n(2n^2+1)}\int\limits_{-\pi}^{+\pi} f(\exp ix)\left(\frac{\sin n\dfrac{y-x}{2}}{\sin\dfrac{y-x}{2}}\right)^4 dy,$$

что и требовалось доказать.

Из доказанного утверждения в силу теоремы Джексона следует неравенство

$$|f(\exp ix)-F_n(\exp ix)|\leqq 6\omega\left(\frac{1}{n}\right)$$

и, следовательно, неравенство (31) действительно имеет место и во всем круге $|z| \leqq 1$.

Остается доказать неравенства (32). Так как

$$\frac{dF_n(\exp ix)}{dx} = i \exp ix\, F_n'(\exp ix),$$

то на окружности $|z| = 1$ имеет место равенство

$$|F_n'(z)| = \left| \frac{dF_n(\exp ix)}{dx} \right|.$$

Отсюда в силу леммы 1 и равенства (33) получаем

$$|F'(z)| \leqq 20\, n\omega\left(\frac{1}{n}\right) = 10(2n)\, \omega\left(\frac{1}{n}\right).$$

Так как это равенство имеет место при $|z| = 1$, то оно должно выполняться и если $|z| < 1$ и, следовательно, (32) имеет место в случае $m = 1$.

В общем случае (32) может быть получено из предыдущего неравенства методом полной индукции, так как по теореме М. Риса для многочленов $P_r(z)$ r-ой степени в круге $|z| \leqq 1$ из неравенства

$$|P_r(z)| \leqq M$$

следует соотношение

$$|P_r'(z)| \leqq Mr.$$

Таким образом

$$|F_n^{(m)}(z)| \leqq 10(2n)\, \check{\omega}\left(\frac{1}{n}\right)(2n-3)(2n-4)\cdots(2n-m-1) \leqq 10(2n)^m\, \omega\left(\frac{1}{n}\right),$$

что и требовалось доказать.

§ 4

В этом §-е, оценив сумму модулей фундаментальных многочленов, мы исследуем сходимость (0, 2)-интерполяционных многочленов.

Лемма 3. В круге $|z| \leqq 1$

(34) $$\sum_{k=1}^{n} |A_k(z)| < 3(3 + \log n),$$

(35) $$\sum_{k=1}^{n} |B_k(z)| < 2\,\frac{3 + \log n}{n(n-1)}.$$

Доказательство. Известно, что для фундаментальных многочленов Лагранжева интерполирования в круге $|z| \leqq 1$ имеет место неравенство

$$\sum_{k=1}^{n} |l_k(z)| < 3 + \log n. \tag{36}$$

Отсюда с помощью неравенства М. Риса легко получается неравенство

$$\sum_{k=1}^{n} |l_k''(z)| < (3 + \log n)\, n^2. \tag{37}$$

Действительно, если $|z_0| \leqq 1$, то

$$|l_k''(z_0)| = \varepsilon_k l_k''(z_0),$$

где $|\varepsilon_k| = 1$, $\sum_{k=1}^{n} \varepsilon_k l_k(z)$ есть многочлен не выше $n-1$-ой степени, следовательно

$$\sum_{k=1}^{n} |l_k''(z_0)| = \sum_{k=1}^{n} \varepsilon_k l_k''(z_0) \leqq (n-1)(n-2) \max_{|z| \leqq 1} \left| \sum_{k=1}^{n} \varepsilon_k l_k(z) \right| <$$

$$< n^2 \max_{|z| \leqq 1} \sum_{k=1}^{n} |l_k(z)| < n^2 (3 + \log n),$$

что и требовалось доказать.

Из (3) при $|z| \leqq 1$ получаем

$$\sum_{k=1}^{n} |A_k(z)| \leqq \sum_{k=1}^{n} |l_k(z)| + \frac{1}{n} \int_0^1 x^{\frac{n+1}{2}} \sum_{k=1}^{n} |l_k''(x \exp i \arg z)|\, dx. \tag{38}$$

Так как

$$\int_0^1 x^{\frac{n+1}{2}}\, dx = \frac{2}{n+1} < \frac{2}{n},$$

то отсюда и из (36)—(38) следует (34):

$$\sum_{k=1}^{n} |A_k(z)| < 3 + \log n + \frac{1}{n} n^2 (3 + \log n) \frac{2}{n} = 3(3 + \log n).$$

Аналогичным образом можно доказать (35), исходя из (4) и (36). Порядок оценок (34) и (35), не может быть улучшен.

Теорема 2. Пусть функция $f(z)$ регулярна при $|z| < 1$ и непрерывна при $|z| \leqq 1$. Обозначим через $\omega(\delta)$ модуль непрерывности функции $f(\exp ix)$. Если

$$\lim_{\delta = +0} \omega(\delta) \log \delta = 0 \tag{39}$$

и

$$\beta_k = o\left(\frac{n^2}{\log n}\right) \qquad (k = 1, 2, \ldots, n), \tag{40}$$

то последовательность многочленов

$$R_n(z)=\sum_{k=1}^{n} f(z_k)\,A_k(z)+\sum_{k=1}^{n}\beta_k B_k(z) \tag{41}$$

в круге $|z|\leqq 1$ равномерно сходится к $f(z)$.

Доказательство. Так как определенная формулой (30) функция $F_n(z)$ есть многочлен не выше $2n-2$-ой степени, то имеет место тождество

$$F_n(z)=\sum_{k=1}^{n} F_n(z_k)\,A_k(z)+\sum_{k=1}^{n} F_n''(z_k)\,B_k(z).$$

Используя это соотношение и (41), получаем

$$f(z)-R_n(z)=f(z)-F_n(z)+F_n(z)-R_n(z)=f(z)-F_n(z)+$$

$$+\sum_{k=1}^{n} F_n(z_k)\,A_k(z)+\sum_{k=1}^{n} F_n''(z_k)\,B_k(z)-\sum_{k=1}^{n} f(z_k)\,A_k(z)-\sum_{k=1}^{n}\beta_k B_k(z)=$$

$$=f(z)-F_n(z)+\sum_{k=1}^{n}[F_n(z_k)-f(z_k)]\,A_k(z)+\sum_{k=1}^{n}[F_n''(z_k)-\beta_k]\,B_k(z).$$

Следовательно,

$$|f(z)-R_n(z)|\leqq|f(z)-F_n(z)|+\sum_{k=1}^{n}|f(z_k)-F_n(z_k)|\,|A_k(z)|+$$

$$+\sum_{k=1}^{n}[|F_n''(z_k)|+|\beta_k|]\,|B_k(z)|.$$

Отсюда, из (31), (32), (34), (35), (40) и (39) получаем

$$|f(z)-R_n(z)|\leqq 6\omega\left(\frac{1}{n}\right)+6\omega\left(\frac{1}{n}\right)3(3+\log n)+\left[40n^2\omega\left(\frac{1}{n}\right)+\right.$$

$$\left.+o\left(\frac{n^2}{\log n}\right)\right]2\,\frac{3+\log n}{n(n-1)}=O\left[\omega\left(\frac{1}{n}\right)\log n\right]+o(1)=o(1),$$

что и требовалось доказать.

§ 5

В этом §-е дается явный вид $(0, 1, \ldots, r-2, r)$-интерполяционных многочленов, доказывается их единственность и исследуется их сходимость в случае $r=3$.

Теорема 3. Если $r\geqq 2$ натуральное число, $n\geqq r$,

$$h_{k,m}(z)\qquad (k=1, 2, \ldots, n;\ m=0, 1, \ldots, r-2)$$

многочлены $(r-1)n-1$-ой степени, удовлетворяющие условиям

$$h^{(p)}_{k,m}(z_j)=\begin{cases}1, & \text{если} \quad j=k,\ p=m,\\ 0, & \text{если} \quad j\neq k \quad \text{или} \quad p\neq m\end{cases} \tag{42}$$

$$(j,k=1,2,\ldots,n;\ m,p=0,1,\ldots,r-2$$

$$A_{k,m}(z)=h_{k,m}(z)-$$

$$-\frac{1}{r!\,n^{r-1}}\left(z^{\frac{n+1}{2}}-z^{-\frac{n-1}{2}}\right)^{r-1}\int\limits_0^z t^{\frac{(r-1)(n-1)}{2}-1}\sum_{j=1}^{n}h^{(r)}_{k,m}(z_j)\,z_j^r\,l_j(t)\,dt, \tag{43}$$

$$B_k(z)=\frac{z_k^r}{r!\,n^{r-1}}\left(z^{\frac{n+1}{2}}-z^{-\frac{n-1}{2}}\right)^{r-1}\int\limits_1^z t^{\frac{(r-1)(n-1)}{2}-1}\,l_k(t)\,dt, \tag{44}$$

$\alpha_{k,m}$ и β_k любые комплексные числа, то функция

$$R_n(z)=\sum_{m=0}^{r-2}\sum_{k=1}^{n}\alpha_{k,m}A_{k,m}(z)+\sum_{k=1}^{n}\beta_k B_k(z) \tag{45}$$

есть единственный многочлен не выше $rn-1$-ой степени, удовлетворяющий условиям

$$R_n^{(m)}(z_k)=\alpha_{k,m} \quad (k=1,2,\ldots,n;\ m=0,1,\ldots,r-2) \tag{46}$$

$$R_n^{(r)}(z_k)=\beta_k \qquad (k=1,2,\ldots,n). \tag{47}$$

Доказательство. Доказательство этой теоремы совершенно аналогично доказательству теоремы 1.

Исходя из формулы

$$l_k(z)=\frac{1}{n}\sum_{\nu=0}^{n-1}z_k^{n-\nu}z^{\nu},$$

убеждаемся в том, что $A_{k,m}(z)$, $B_k(z)$ и, следовательно, $R_n(z)$ суть многочлены не выше $rn-1$-ой степени.

Выполнение условий (46) следует из (42)—(45), (10) и (12).

Чтобы доказать (47), вычислим $A^{(r)}_{k,m}(z_\nu)$. Вычисляя r-ую производную произведения

$$\omega(z)\int\limits_0^z t^{\frac{(r-1)(n-1)}{2}}\sum_{j=1}^{n}h^{(r)}_{k,m}(z_j)z_j^r\,l_j(t)\,dt$$

в узле z_ν с помощью правила Лейбница, убеждаемся в том, что она равна

$$r\left[\omega(z)^{r-1}\right]^{(r-1)}_{z_\nu}z_\nu^{\frac{(r-1)(n-1)}{2}-1}h^{(r)}_{k,m}(z_\nu)\,z_\nu^r,$$

так как в остальных слагаемых встречаются $\omega(z_\nu)$ или $\omega''(z_\nu)$, равные, как мы видели, нулю, и

$$l_j(z_\nu)=\begin{cases}1, & \text{если } j=\nu\\ 0, & \text{если } j\neq\nu\end{cases}\qquad (j,\nu=1,2,\ldots,n).$$

Но

$$\left[\omega(z)^{r-1}\right]_{z_\nu}^{(r-1)}=(r-1)!\,\omega'(z_\nu)^{r-1}=(r-1)!\,n^{r-1}z_\nu^{-\frac{(r-1)(n+1)}{2}}$$

и, следовательно, наша производная равна

$$r!\,n^{r-1}h_{k,m}^{(r)}(z_\nu),$$

откуда в силу (43) получаем

$$A_{k,m}^{(r)}(z_\nu)=0\qquad (k,\nu=1,2,\ldots,n;\ m=0,1,\ldots,r-2).$$

Аналогичным образом можно убедиться в том, что

$$B_k^{(r)}(z_\nu)=\begin{cases}1, & \text{если } \nu=k\\ 0, & \text{если } \nu\neq k\end{cases}\qquad (\nu,k=1,2,\ldots,n).$$

Из двух последних соотношений и (45) и следует (47).

Из существования интерполяционного многочлена, также как и в случае $r=2$, следует его единственность.

Заметим, что в случае $r=3$

$$\begin{aligned} h_{k,0}(z)&=l_k^2(z)\left[1-\frac{n}{z_k}(z-z_k)\right],\\ h_{k,1}(z)&=l_k(z)(z-z_k)\end{aligned}\qquad (k=1,2,\ldots,n). \tag{48}$$

Наконец, исследуем сходимость (0, 1, 3)-интерполяционных многочленов. Аналогичную теорему, вероятно, можно доказать и для общего случая.

Теорема 4. Пусть функция $f(z)$ непрерывна при $|z|\leqq 1$ и регулярна при $|z|<1$, $\omega(\delta)$ есть модуль непрерывности функции $f(\exp ix)$,

$$\lim_{\delta=+0}\omega(\delta)\log^2\delta=0, \tag{49}$$

$$\alpha_k=o\left(\frac{n}{\log^2 n}\right), \tag{50}$$

$$\beta_k=o\left(\frac{n^3}{\log n}\right)\qquad (k=1,2,\ldots,n). \tag{51}$$

Тогда последовательность многочленов

$$R_n(z)=\sum_{k=1}^{n}f(z_k)A_{k,0}(z)+\sum_{k=1}^{n}\alpha_kA_{k,1}(z)+\sum_{k=1}^{n}\beta_kB_k(z) \tag{52}$$

равномерно сходится в круге $|z|\leqq 1$ к $f(z)$.

Доказательство. Легко видеть, что

$$\sum_{k=1}^{n} |h_{k,m}(z)| = O\left(\frac{\log n}{n^m}\right) \qquad (m = 0, 1). \tag{53}$$

Отсюда нетрудно вывести соотношение

$$\sum_{k=1}^{n} |h_{k,m}^{(3)}(z)| = O(n^{3-m} \log n).$$

Поэтому

$$\sum_{k=1}^{n} \left| \sum_{j=1}^{n} h_{k,m}^{(3)}(z_j)\, z_j^{\nu}\, l_j(t) \right| \leqq \sum_{j=1}^{n} |l_j(t)| \sum_{k=1}^{n} |h_{k,m}^{(3)}(z_j)| =$$

$$= O(\log n)\, O(n^{3-m} \log n) = O(n^{3-m} \log^2 n).$$

Отсюда, из (53) и (43) получаем

$$\sum_{k=1}^{n} |A_{k,m}(z)| = O\left(\frac{\log^2 n}{n^m}\right) \qquad (m = 0, 1). \tag{54}$$

Очевидно

$$\sum_{k=1}^{n} |B_k(z)| = O\left(\frac{\log n}{n^3}\right). \tag{55}$$

Так как

$$F_n(z) = \sum_{m=0}^{1} \sum_{k=1}^{n} F_n^{(m)}(z_k)\, A_{k,m}(z) + \sum_{k=1}^{n} F_n^{(3)}(z_k)\, B_k(z),$$

то в силу (52), (31), (54), (32), (55) и (49)

$$f(z) - R_n(z) = f(z) - F_n(z) + F_n(z) - R_n(z) = f(z) - F_n(z) +$$

$$+ \sum_{m=0}^{1} \sum_{k=1}^{n} F_n^{(m)}(z_k)\, A_{k,m}(z) + \sum_{k=1}^{n} F_n^{(3)}(z_k)\, B_k(z) - \sum_{k=1}^{n} f(z_k)\, A_{k,0}(z) - \sum_{k=1}^{n} \alpha_k A_{k,1}(z) -$$

$$- \sum_{k=1}^{n} \beta_k B_k(z) = f(z) - F_n(z) + \sum_{k=1}^{n} [F_n(z_k) - f(z_k)]\, A_{k,0}(z) +$$

$$+ \sum_{k=0}^{n} [F_n'(z) - \alpha_k]\, A_{k,1}(z) + \sum_{k=1}^{n} [F_n^{(3)}(z_k) - \beta_k]\, B_k(z) = O\left[\omega\left(\frac{1}{n}\right)\right] +$$

$$+ O\left[\omega\left(\frac{1}{n}\right)\right] O(\log^2 n) + \left\{ O\left[n\,\omega\left(\frac{1}{n}\right) + o\left(\frac{n}{\log^2 n}\right)\right]\right\} O\left(\frac{\log^2 n}{n}\right) +$$

$$+ \left\{ O\left[n^3 \omega\left(\frac{1}{n}\right)\right] + o\left(\frac{n^3}{\log n}\right)\right\} O\left(\frac{\log n}{n^3}\right) = o(1),$$

что и требовалось доказать.

(Поступило 13. VII. 1959.)

ÜBER DEN AFFINZUSAMMENHANG VON, ZU TANGENTIALRÄUMEN GEHÖRENDEN PRODUKTRÄUMEN

Von

L. TAMÁSSY (Debrecen)

(*Vorgelegt von* O. VARGA)

Einleitung

Der Begriff des affinzusammenhängenden Raumes wurde von H. WEYL eingeführt [5]. WEYL nennt den Raum dann affinzusammenhängend, wenn die Vektoren der in zwei beliebigen benachbarten Punkten P und P' genommenen Vektorräumen durch eine affine Abbildung miteinander verkünpft sind. Die Parallelverschiebung, d. h. die Zueinanderordnung der in benachbarten Punkten definierten Tensoren geschieht schon mit Hilfe der parallel verschobenen Vektoren, unter Anwendung der Zerspaltung des Tensors auf Komponenten nach den Vektoren eines adjungierten ko- und kontravarianten n-Beins.

Die Gesamtheit der Tensoren zweiter Ordnung in den Punkten P bzw. P' bilden zwei Vektorräume von n^2 Dimension, genauer: sie bilden die Produkträume von n^2 Dimension der n-dimensionalen Tangentialräume. Den Vorangehenden gemäß sind diese Vektorräume einander affin zugeordnet. Aber diese Zuordnung geschieht nicht unmittelbar, sondern mit Hilfe der parallel verschobenen Vektoren. Es ergibt sich die Frage, ob in dieser Weise jede affine Zuordnung dieser Tensoren entstehen kann, oder nur gewisse von diesen. Im letzteren Falle ist es nämlich zu erwarten, daß man durch unmittelbare Zuordnung der Tensoren zweiter Stufe für diese einen allgemeineren affinen Zusamennhang definieren kann. Wir werden zeigen, daß diese letzt erwähnte Möglichkeit wirklich besteht.

Wir können mit Hilfe des von WEYL eingeführten Zusammenhanges zwischen den in den benachbarten Punkten P bzw. P' erklärten sämtlichen Tensoren eine Verbindung zustande bringen, aber diese Verbindung führt auch aus der engeren Menge der Invarianten und Tensoren, die in ihren kontra- bzw. kovarianten Indizes von gerader Anzahl sind, nicht hinaus. Wir werden zwischen diesen Mengen einen Zusammenhangtyp angeben, welcher zum Vorangehenden ähnliche Eigenschaften besitzt, aber allgemeiner ist als jener Zusammenhang, und über dieser engeren Menge denselben als einen Spezialfall enthält.

Für die Zwecke der Definition des affinen Zusammenhanges über der obenerwähnten engeren Menge scheinen die von WEYL gebrauchten zwei

Bedingungen axiomatischen Charakters nicht leicht anwendbar zu sein, weil der Begriff des dort vorkommenden infinitesimalen Parallelogrammes mit der Parallelverschiebung der Vektoren in enger Verbindung steht, und somit hier gar nicht erklärt werden kann. Aber beim Zustandebringen des affinen Zusammenhanges ist ein gewisse Forderungen befriedigender absoluter Differentialoperator sehr vorteilhaft zu benutzen. Einen solchen benutzt O. VARGA in seiner den affinen Zusammenhang der Linienelementenräume einleitenden Arbeit [4].

So wird vor allem über der erwähnten Menge der Tensoren von gerader Stufe die allgemeinste Form jenes absoluten Differentialoperators bestimmt, welcher denselben Forderungen genügt wie der gewöhnliche. Wir werden den affinen Zusammenhang durch das Verschwinden dieses neuen absoluten Differentials erklären. Dies führen wir im § 1 aus. Wir bemerken hier, daß das für die Tensoren zweiter Stufe sich hier als Ergebnis ergebende absolute Differential auch schon bei E. BOMPIANI [1] vorkommt, ohne den Beweis, daß dieses den allgemeinsten Typ von gewissen natürlichen Forderungen genügenden Operationen darstellt. Im § 2 zeigen wir, wie dieser neue Zusammenhang den Alten als Spezialfall enthält, und wir bestimmen Kriterien für das Auftreten des Falles der Reduktion. Im § 3 stellen wir eine notwendige und eine hinreichende Bedingung für die Äquivalenz zweier solcher Räume auf. Im § 4 bestimmen wir die Krümmungen und die Torsion unseres Raumes, und zeigen auf die geometrische Deutung dieser Tensoren. Im § 5 charakterisieren wir den Fall der Reduktion auf rein geometrischem Wege.

Zum Schluß bemerken wir, daß aus den folgenden klar hervorgeht, daß die Rolle der Menge der Tensoren gerader Stufe durch eine beliebige Menge übernommen werden kann, welche aus solchen Tensoren besteht, die in ihren ko- bzw. kontravarianten Indizes von nk-ter Ordnung sind, wo $k>2$ eine fixierte und n eine beliebige natürliche Zahl ist. In diesem Falle ist aber die Gestalt der Zusammenhänge und die Fassung der Ergebnisse offensichtlich komplizierter. Wir begnügen uns mit der hier angeführten Bemerkung dieser Tatsache und werden dies in unseren Ergebnissen nicht zum Ausdruck bringen.

§ 1. Übertragungskoeffizienten

Sei $\mathfrak{P}$ die Menge, die aus den Invarianten und aus denjenigen Tensoren des n-dimensionalen Raumes besteht, die in ihren ko- bzw. kontravarianten Indizes von gerader Ordnung sind. Wir suchen über $\mathfrak{P}$ einen solchen im allgemeinen vom Ort abhängigen absoluten Differentialoperator, der in den Komponenten des Tensors, in den ersten Ableitungen der Komponenten des

Tensors, und in den Koordinatendifferentialen (in jedem einzelnen derselben) linear ist, einem Tensor einen Tensor gleicher Stufe ordnet, unabhängig vom Koordinatensystem ist, für Summe und Produkt die gewöhnlichen Differentiationsregeln bestehen, und einer Invariante das gewöhliche Differential zuordnet. D. h. für diesen Operator ϑ gilt

$$\vartheta T = L_i\left(T, \frac{\partial T}{\partial x}, x\right) dx^i, \tag{1}$$

$$\overline{\vartheta T}^{i\ldots}{}_{\ldots s} = \frac{\partial \overline{x}^i}{\partial x^k} \cdots \frac{\partial x^l}{\partial \overline{x}^s} \vartheta T^{k\ldots}{}_{\ldots l}, \tag{2}$$

$$\overline{\vartheta T} = \vartheta \overline{T}, \tag{3}$$

$$\vartheta(T+U) = \vartheta T + \vartheta U, \tag{4}$$

$$\vartheta(T \cdot U) = \vartheta T \cdot U + T \cdot \vartheta U, \tag{5}$$

$$\vartheta I = dI, \tag{6}$$

wo T und U Elemente von $\mathfrak{P}$ sind, L_i in T und $\frac{\partial T}{\partial x}$ bilinear ist, I eine Invariante und d das gewöhnliche Differential bezeichnet. Wir haben mit $\overline{x}$ die neue Koordinaten und mit "$-$" im allgemeinen die im neuen Koordinatensystem genommenen Größen bezeichnet.

Unsere erste Aufgabe wird die Bestimmung der Form des allgemeinsten solchen Operators sein. Wir beschränken uns zuerst nur auf Tensoren zweiter Stufe. Aus den hier gewonnenen Ergebnissen werden wir die allgemeine Form schon leicht bestimmen können.

ϑT^{ij} kann wegen (1) und (4) in T^{kl} und in $\frac{\partial T^{kl}}{\partial x^r}$ zusammen nur linear sein, und im Ausdruck von ϑT^{ij} kann kein von T^{kl} und $\frac{\partial T^{kl}}{\partial x^r}$ freies, nur von x abhängiges Glied vorkommen. Daher ist

$$\vartheta T^{ij}(x) = \psi^{ij}{}_{kl}{}^{r}{}_{t}(x) \frac{\partial T^{kl}}{\partial x^r} dx^t + \gamma_{kl}{}^{ij}{}_{t}(x) T^{kl} dx^t. \tag{7}$$

Damit ϑT^{ij} vom Koordinatensystem unabhängig sei, und einen kontravarianten Tensor zweiter Stufe bildet, müssen sich ψ und γ nach

$$\overline{\psi}^{ab}{}_{cd}{}^{e}{}_{f} = \frac{\partial \overline{x}^a}{\partial x^i} \frac{\partial \overline{x}^b}{\partial x^j} \frac{\partial x^k}{\partial \overline{x}^c} \frac{\partial x^l}{\partial \overline{x}^d} \frac{\partial \overline{x}^e}{\partial x^r} \frac{\partial x^u}{\partial \overline{x}^f} \psi^{ij}{}_{kl}{}^{r}{}_{u}, \tag{8}$$

$$\begin{aligned}\overline{\gamma}_{cd}{}^{ab}{}_{f} = {} & \frac{\partial \overline{x}^a}{\partial x^i} \frac{\partial \overline{x}^b}{\partial x^j} \frac{\partial x^k}{\partial \overline{x}^c} \frac{\partial x^l}{\partial \overline{x}^d} \frac{\partial x^t}{\partial \overline{x}^f} \gamma_{kl}{}^{ij}{}_{t} + \\ & + \frac{\partial \overline{x}^a}{\partial x^i} \frac{\partial \overline{x}^b}{\partial x^j} \left(\frac{\partial^2 x^k}{\partial \overline{x}^c \partial \overline{x}^m} \frac{\partial x^l}{\partial \overline{x}^d} + \frac{\partial x^k}{\partial \overline{x}^c} \frac{\partial^2 x^l}{\partial \overline{x}^d \partial \overline{x}^m} \right) \frac{\partial \overline{x}^m}{\partial x^r} \frac{\partial x^t}{\partial \overline{x}^f} \psi^{ij}{}_{kl}{}^{r}{}_{t}\end{aligned} \tag{9}$$

transformieren.

Betrachten wir jetzt den Ausdruck $T^{ij}(x)U_{ij}(x)=I(x)$. Das absolute Differential von $I(x)$ ist gemäß (5) und auch gemäß (6) berechenbar. Durch Vergleich ergibt sich

$$(10)\qquad T^{ij}\vartheta U_{ij}=T^{ij}(dU_{ij}-\gamma_{ij}{}^{kl}{}_{t}U_{kl}dx^{t})+U_{ij}\left(\frac{\partial T^{ij}}{\partial x^{t}}-\psi^{ij}{}_{kl}{}^{r}{}_{t}\frac{\partial T^{kl}}{\partial x^{r}}\right)dx^{t}.$$

Dies muß für einen beliebigen Tensor T^{ij} und für beliebiges dx^t gelten. So z. B. auch wenn jede Komponente von T^{ij} außer $T^{11}=1$ verschwindet. Daraus ist es ersichtlich, daß für die Existenz eines von $\dfrac{\partial T^{kl}}{\partial x^{r}}$ unabhängigen, sowohl (5) als auch (6) befriedigenden ϑT^{ij} die Bedingung

$$\psi^{ij}{}_{kl}{}^{r}{}_{t}=\delta^{i}_{k}\delta^{j}_{l}\delta^{r}_{t}$$

notwendig ist. Eine solche Wahl von ψ ist vom Koordinatensystem unabhängig, da nach (8) auch ψ ein Tensor ist. So folgt aus (7), (10) und (9)

$$(11)\qquad \vartheta T^{ij}=dT^{ij}+\gamma_{kl}{}^{ij}{}_{t}T^{kl}dx^{t},$$

$$(12)\qquad \vartheta U_{ij}=dU_{ij}-\gamma_{ij}{}^{kl}{}_{t}U_{kl}dx^{t},$$

$$(13)\qquad \bar{\gamma}_{cd}{}^{ab}{}_{f}=\frac{\partial\bar{x}^{a}}{\partial x^{m}}\frac{\partial^{2}x^{m}}{\partial\bar{x}^{c}\partial\bar{x}^{f}}\delta^{b}_{d}+\delta^{a}_{c}\frac{\partial\bar{x}^{b}}{\partial x^{m}}\frac{\partial^{2}x^{m}}{\partial\bar{x}^{d}\partial\bar{x}^{f}}+\frac{\partial\bar{x}^{a}}{\partial x^{i}}\frac{\partial\bar{x}^{b}}{\partial x^{j}}\frac{\partial x^{k}}{\partial\bar{x}^{c}}\frac{\partial x^{l}}{\partial\bar{x}^{d}}\frac{\partial x^{t}}{\partial\bar{x}^{f}}\gamma_{kl}{}^{ij}{}_{t}.$$

Es ist mit Rücksicht auf (13) leicht nachzuprüfen, daß die entsprechend (12) gewonnenen Werte von ϑU_{ij} auch die Forderungen (1)—(4) befriedigen.

Wollen wir jetzt $\vartheta A^{ij\ldots}{}_{op\ldots}$ bestimmen, wo A ein Element von höherer Ordnung von $\mathfrak{P}$ ist, so überschieben wir A mit Elementen von zweiter Ordnung von $\mathfrak{P}$, solange bis wir eine Invariante bekommen. Darauf wenden wir (5) und (6) an, dann folgt durch Vergleich

$$(14)\qquad \vartheta A^{ij\ldots}{}_{op\ldots}=dA^{ij\ldots}{}_{op\ldots}+(\gamma_{kl}{}^{ij}{}_{t}A^{kl\ldots}{}_{op\ldots}+\cdots-\gamma_{op}{}^{kl}{}_{t}A^{ij\ldots}{}_{kl\ldots}-\cdots)dx^{t}.$$

Das Erfülltsein von (1)—(6) läßt sich wieder gleich überprüfen. Mit Rücksicht auf den Gang der Überlegungen dieses Paragraphen können wir behaupten

SATZ 1. *Im Kreis der Tensoren, die in ihren kontra- bzw. kovarianten Indizes von gerader Ordnung sind, ist* (14) *der absolute Differentialoperator von allgemeinster Form, der die Bedingungen* (1)—(6) *befriedigt.* (13) *ist das Transformationsgesetz der den Operator festsetzenden Größen.*

Hiernach definieren wir die Parallelität der Elemente von $\mathfrak{P}$ in zwei benachbarten Punkten $P(x)$ und $P'(x+dx)$.

DEFINITION. $A(P)$ ist parallel verschoben von P nach P', wenn $\vartheta A=0$ ist.

Wir bemerken: a) Das Transformationsgesetz (13) erfüllt die Eigenschaft der Transitivität. Daher bilden die γ bezüglich der stetig differenzierbaren und umkehrbaren Transformationen $x^i = x^i(\bar{x})$ eine Transformationsgruppe. — b) Es hat keinen Sinn, von der Symmetrie von γ in zwei seiner oberen bzw. unteren Indizes zu sprechen, weil diese Eigenschaften bei einer Transformation im allgemeinen verloren gehen. Aber die Relation $\gamma_{kl}{}^{ij}{}_t = \gamma_{lk}{}^{ji}{}_t$ bleibt invariant. Bezüglich der Eigenschaft, daß zusammen mit einem symmetrischen (bzw. antisymmetrischen) Tensor T^{ij} auch ϑT^{ij} symmetrisch (bzw. antisymmetrisch) sei, ist eben die erwähnte Relation notwendig und hinreichend.

§ 2. Der Fall der Reduktion

1. Den Relationen (§ 1, (1)—(6)) genügt über $\mathfrak{P}$ auch

$$DA^{ij\dots}{}_{op\dots} = dA^{ij\dots}{}_{op\dots} + (\Gamma_t{}^i{}_k A^{kj\dots}{}_{op\dots} + \cdots - \Gamma_t{}^k{}_o A^{ij\dots}{}_{kp\dots} - \cdots)dx^t, \tag{1}$$

$$\Gamma_b{}^a{}_c = \frac{\partial \bar{x}^a}{\partial x^m} \frac{\partial^2 x^m}{\partial \bar{x}^b \partial \bar{x}^c} + \frac{\partial \bar{x}^a}{\partial x^i} \frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial x^k}{\partial \bar{x}^c} \Gamma_j{}^i{}_k. \tag{2}$$

Daher ist es möglich, einen affinen Zusammenhang auch durch das Verschwinden von DA zu definieren. Bezeichnen wir den im vorigen Paragraphen über $\mathfrak{P}$ mit Hilfe von γ definierten affinzusammenhängenden Raum mit $\{\mathfrak{P}, \gamma\}$, so können wir diesen Raum mit $\{\mathfrak{P}, \Gamma\}$ bezeichnen. Wir zeigen, daß $\{\mathfrak{P}, \gamma\}$ den $\{\mathfrak{P}, \Gamma\}$ als Spezialfall enthält. Nach Satz 1 wäre es hinreichend nur darauf hinweisen, daß ϑ mit D über $\mathfrak{P}$ nicht äquivalent ist. Wir wollen aber auch die Frage untersuchen, wie und wann ϑ den Operator D enthält. Der Einfachheit wegen beschränken wir uns wieder nur auf Tensoren zweiter Stufe, dies gibt aber schon eine klare Übersicht über den allgemeinen Fall.

Es seien im Koordinatensystem x die Γ und damit auch die

$$DT^{ij} = dT^{ij} + (\Gamma_k{}^i{}_t T^{kj} + \Gamma_l{}^j{}_t T^{il})dx^t \tag{3}$$

gegeben. (§ 1, (11)) und (3) unterscheiden sich nur in den Koeffizienten der entsprechenden T^{kl}. Definieren wir also jetzt die γ so, daß die Koeffizienten der T^{kl} in den zwei Ausdrücken einander gleich seien, dann ist

$$\gamma_{kl}{}^{ij}{}_t = \delta_l^j \Gamma_k{}^i{}_t + \delta_k^i \Gamma_l{}^j{}_t. \tag{4}$$

Dies ist in einer auf Null reduzierten Form mit Rücksicht auf (2) und (§ 1, (13)) eine Tensorrelation. — Weiter kann (4) bei gegebenen γ im allgemeinen nicht gelten, doch müssen gemäß (4) z. B. von den Größen $\gamma_{kl}{}^{ij}{}_t$ mehrere verschwinden. Es gilt daher der

Satz 2. *Der mit Hilfe des* ϑ *definierte affine Zusammenhang enthält den über* $\mathfrak{P}$ *mit Hilfe von D definierten affinen Zusammenhang als Spezialfall. Damit sich der erstere auf den letzteren reduziere, ist das Bestehen von* (4) *notwendig und hinreichend.*

2. Da die Lösbarkeit des Gleichungssystems (4) nur von den γ abhängt, und von dem Koordinatensystem unabhängig ist, so müssen zwischen den γ von dem Koordinatensystem unabhängige Relationen existieren, die notwendige und hinreichende Bedingungen der Lösbarkeit ausdrücken. Wir wollen diese Bedingungen jetzt bestimmen.[1] Dabei werden sich die unabhängigen Gleichungen in (4) herausstellen, die alle nur einen einzigen Unbekannten enthalten, und so werden wir für die Lösung von (4) — im Falle der Lösbarkeit — ein sehr einfaches Verfahren erhalten. Wir werden diese Bedingungen erst in einem beliebig gewählten Koordinatensystem bestimmen, und darauf werden wir zeigen, daß sie koordinateninvarianten Charakter haben.

Da die rechte Seite von (4) sowohl in den Indizes i und j als auch in den Indizes k und l symmetrisch ist, so ist zur Lösbarkeit

$$\gamma_{kl}{}^{ij}{}_t - \gamma_{lk}{}^{ji}{}_t = 0 \tag{5}$$

notwendig. Weiterhin ist, wie aus (4) ersichtlich, auch das Bestehen der Relation

$$\gamma_{kl}{}^{ij}{}_t = 0 \quad \text{wenn} \quad i \neq k \quad \text{und} \quad j \neq l \tag{6}$$

notwendig. Im folgenden setzen wir das Bestehen von (5) und (6) voraus.

Jetzt zerspalten wir die Gleichungen des Gleichungssystems (4) in drei Teile:

A) Es ist $i \neq k$ und $j \neq l$.

Diese Gleichungen können wir vernachlässigen, da sie mit Rücksicht auf (6) für jedes $\Gamma_p{}^{o}{}_t$ erfüllt sind.

B) Es ist $i = k$, aber $j \neq l$.

Somit ist

$$\gamma_{1l}{}^{1j}{}_t = \Gamma_l{}^j{}_t, \ldots, \gamma_{nl}{}^{nj}{}_t = \Gamma_l{}^j{}_t \qquad (j \neq l). \tag{7}$$

Die notwendige und hinreichende Bedingung der Lösbarkeit dieses Teilsystems ist

$$\gamma_{r^*l}{}^{r^*j}{}_t - \gamma_{s^*l}{}^{s^*j}{}_t = 0 \qquad (j \neq l). \tag{8}$$

[1] Die einfachste Bedingung solcher Art ist die, welche die Übereinstimmung der Ränge der aus den Koeffizienten des linearen Gleichungssystems (4) gebildeten Matrix und der erweiterten Matrix ausdrückt. Dies ist aber eine verhältnismäßig komplizierte Relation, und wir werden eine durchsichtigere angeben.

(Wenn die Summation für zwei Indizes vorgenommen werden müßte, aber diese mit Stern versehen sind, so bleibt die Summation weg.)

C) Es ist $i=k$ und $j=l$.

Die hierher gehörenden Gleichungen sind

$$(9)\quad \begin{aligned} &\text{a)}\quad \gamma_{11}{}^{11}{}_t = 2\Gamma_1{}^1{}_t, \ldots, \gamma_{nn}{}^{nn}{}_t = 2\Gamma_n{}^n{}_t, \\ &\text{b)}\quad \gamma_{12}{}^{12}{}_t = \Gamma_1{}^1{}_t + \Gamma_2{}^2{}_t, \ldots, \gamma_{n-1\,n}{}^{n-1\,n}{}_t = \Gamma_{n-1}{}^{n-1}{}_t + \Gamma_n{}^n{}_t. \end{aligned}$$

Die notwendige und hinreichende Bedingung der Lösbarkeit dieses Teilsystems ist

$$(10)\qquad \gamma_{k^*l^*}{}^{k^*l^*}{}_t - \frac{1}{2}\left(\gamma_{k^*k^*}{}^{k^*k^*}{}_t + \gamma_{l^*l^*}{}^{l^*l^*}{}_t\right) = 0.$$

Aus (§ 1, (13)) ist es gleich ersichtlich, daß (5) und (6) Tensorrelationen sind. Daß auch (8) eine Tensorrelation bildet, ist folgendermaßen einzusehen: Wegen des Wegfallens der zweiten Ableitungen ist

$$\bar{\gamma}_{a^*d}{}^{a^*b}{}_f - \bar{\gamma}_{c^*d}{}^{c^*b}{}_f = \frac{\partial \bar{x}^{a^*}}{\partial x^i}\frac{\partial \bar{x}^b}{\partial x^j}\frac{\partial x^k}{\partial \bar{x}^{a^*}}\frac{\partial x^l}{\partial \bar{x}^d}\frac{\partial x^t}{\partial \bar{x}^f}\gamma_{kl}{}^{ij}{}_t -$$
$$- \frac{\partial \bar{x}^{c^*}}{\partial x^o}\frac{\partial \bar{x}^b}{\partial x^j}\frac{\partial x^q}{\partial \bar{x}^{c^*}}\frac{\partial x^l}{\partial \bar{x}^d}\frac{\partial x^r}{\partial \bar{x}^f}\gamma_{ql}{}^{oj}{}_r \qquad (d \neq b).$$

Im ersten Glied auf der rechten Seite fallen wegen (6) die Glieder mit $i \neq k$ und $j \neq l$ weg. Der im Falle $i=k$ und $j \neq l$ aus diesem Glied hervorgehende Ausdruck ist

$$\sum_{i^*=1}^{n} \frac{\partial \bar{x}^{a^*}}{\partial x^{i^*}}\frac{\partial x^{i^*}}{\partial \bar{x}^{a^*}}\frac{\partial \bar{x}^b}{\partial x^j}\frac{\partial x^l}{\partial \bar{x}^d}\frac{\partial x^t}{\partial \bar{x}^f}\gamma_{i^*l}{}^{i^*j}{}_t.$$

Wegen (7) ist aber jedes in diesem Ausdruck vorkommende γ gleich $\gamma_{k^*l}{}^{k^*j}{}_t$, wo k^* ein beliebiger festgehaltener Index ist. So ist dieser Ausdruck gleich

$$\frac{\partial \bar{x}^b}{\partial x^j}\frac{\partial x^l}{\partial \bar{x}^d}\frac{\partial x^t}{\partial \bar{x}^f}\gamma_{k^*l}{}^{k^*j}{}_t.$$

Im Falle $j=l$ (sowohl für $i \neq k$ als auch für $i=k$) erhalten wir wegen den auftretenden δ_d^b neuestens den Wert Null. Durch ein ähnliches Verfahren erhalten wir im zweiten Glied der rechten Seite

$$\bar{\gamma}_{a^*d}{}^{a^*b}{}_f - \bar{\gamma}_{c^*d}{}^{c^*b}{}_f = \frac{\partial \bar{x}^b}{\partial x^j}\frac{\partial x^l}{\partial \bar{x}^d}\frac{\partial x^t}{\partial \bar{x}^f}\left(\gamma_{k^*l}{}^{k^*j}{}_t - \gamma_{r^*l}{}^{r^*j}{}_t\right).$$

Daß auch (10) unabhängig von Koordinatensystem ist, läßt sich ähnlich beweisen. Damit haben wir den Beweis erbracht für den folgenden

Satz 3. *Die von Koordinatensystem unabhängigen notwendigen und hinreichenden Bedingungen für die Lösbarkeit von* (4), *d. h. die notwendigen und hinreichenden Bedingungen dafür, daß* $\{\mathfrak{P}, \gamma\}$ *sich auf einem* $\{\mathfrak{P}, \Gamma\}$ *reduziere, sind* (5), (6), (8) *und* (10). *Die Werte der* $\Gamma_j{}^i{}_k$ *liefern in diesem Falle* (7) *und* (9a).

Aus (7) und (9) ist es ersichtlich, daß die Lösung von (4) eindeutig ist. D. h. es gilt der

Satz 4. *Ein* $\{\mathfrak{P}, \gamma\}$ *reduziert sich höchstens auf eine Weise auf einen* $\{\mathfrak{P}, \Gamma\}$.

§ 3. Die Äquivalenz

Ein auf das Koordinatensystem x bezogener Raum $\{\mathfrak{P}, \gamma\}$ und ein auf das Koordinatensystem $\bar{x}$ bezogener Raum $\{\mathfrak{P}, \bar{\gamma}\}$ sind miteinander dann äquivalent, wenn es ein solches stetig differenzierbares, eindeutig umkehrbares Funktionensystem

$$x^i = x^i(\bar{x}) \tag{1}$$

gibt, welches die Punkte der zwei Räume einander so zuordnet, daß zwischen den in den einander entsprechenden Punkten genommenen γ und $\bar{\gamma}$ die Relation (§ 1, (13)) gilt. Existiert ein die erwähnten Eigenschaften besitzendes Funktionensystem (1), so ist (§ 1, (13)) mit

$$\begin{gathered}\bar{\gamma}_{cd}{}^{ab}{}_f(\bar{x})\, p^g_a p^h_b = p_c{}^g{}_f\, p^h_d + p_d{}^h{}_f\, p^g_c + \gamma_{kl}{}^{gh}{}_t(x)\, p^k_c p^l_d p^t_f \\ \left(p^g_a(\bar{x}) = \frac{\partial x^g}{\partial \bar{x}^a}; \quad p_c{}^g{}_f(\bar{x}) = \frac{\partial^2 x^g}{\partial \bar{x}^c \partial \bar{x}^f}\right)\end{gathered} \tag{2}$$

äquivalent, was für gegebene γ und $\bar{\gamma}$ ein für das Funktionensystem (1) aufgestelltes Differentialgleichungssystem ist. Wir nehmen eine Umgestaltung von (2) in ein sogenanntes gemischtes System vor, und werden mit Hilfe desselben für die Existenz einer die erwünschten Eigenschaften besitzenden Lösung eine notwendige und eine hinreichende Bedingung ableiten.

Nehmen wir die Existenz einer den obenerwähnten Eigenschaften verfügenden Lösung an, und bezeichnen wir die partiellen Ableitungen der Inversen $\bar{x}^j = \bar{x}^j(x)$ nach x^k von (1) mit q^j_k. Aus (2) ergibt sich

$$\begin{gathered}p^g_a p^h_b(\bar{\gamma}_{fc}{}^{ba}{}_d + \bar{\gamma}_{df}{}^{ba}{}_c - \bar{\gamma}_{cd}{}^{ab}{}_f) = p_c{}^g{}_d\, p^h_f + p_d{}^h{}_c\, p^g_f + \\ + (\gamma_{tk}{}^{hg}{}_l + \gamma_{lt}{}^{hg}{}_k - \gamma_{kl}{}^{gh}{}_t)\, p^k_c p^l_d p^t_f .\end{gathered}$$

Überschiebung mit q^f_g ergibt

$$p_c{}^h{}_d = \frac{1}{n+1}\left(\bar{G}_c{}^b{}_d\, p^h_b - G_k{}^h{}_l\, p^k_c\, p^l_d\right), \tag{3}$$

wo

$$G_k{}^h{}_t \equiv \gamma_{rk}{}^{kr}{}_t + \gamma_{tr}{}^{hr}{}_k - \gamma_{kt}{}^{rh}{}_r \tag{4}$$

ist. Das Transformationsgesetz von G ist

$$\bar{G}_j{}^s{}_i = \frac{\partial \bar{x}^s}{\partial x^m}\left|(n+1)\frac{\partial^2 x^m}{\partial \bar{x}^i \partial \bar{x}^j} + \frac{\partial x^v}{\partial \bar{x}^j}\frac{\partial x^t}{\partial \bar{x}^i} G_v{}^m{}_t\right|. \tag{5}$$

Unser gemischtes System von Typ

$$\text{(6)}\quad \begin{aligned} &\text{a)}\quad \frac{\partial \Theta^\alpha}{\partial \bar{x}^i} = \psi^\alpha_i(\Theta, \bar{x}) \qquad (\alpha = 1, 2, \ldots, M),\\ &\text{b)}\quad F_0^{\lambda_0}(\Theta, \bar{x}) = 0 \qquad (\lambda_0 = 1, 2, \ldots, S_0) \end{aligned}$$

besteht aus den Differentialgleichungen

$$\text{(7)}\quad \begin{aligned} &\text{a)}\quad \frac{\partial x^i}{\partial \bar{x}^a} = p^i_a(\bar{x}),\\ &\text{b)}\quad \frac{\partial p^g_c}{\partial \bar{x}^m} = \frac{1}{n+1}\left[\bar{G}_c{}^a{}_d(\bar{x})\, p^g_a - G_k{}^g{}_l(x)\, p^k_c\, p^l_d\right],\\ &\text{c)}\quad \frac{\partial p_c{}^g{}_d}{\partial \bar{x}^m} = \frac{1}{n+1}\Bigg[\frac{\partial \bar{G}_c{}^b{}_d}{\partial \bar{x}^m} p^g_b + \bar{G}_c{}^b{}_d\, p_b{}^g{}_m - \frac{\partial G_i{}^g{}_j}{\partial x^u} p^u_m\, p^i_c\, p^j_d -\\ &\qquad\qquad - G_i{}^g{}_j\left(p_c{}^i{}_m\, p^j_d + p^i_c\, p_d{}^j{}_m\right)\Bigg], \end{aligned}$$

und aus der Skalarrelation (2). Hier wird die Rolle der Θ^α von den x^i, p^g_c und $p_c{}^g{}_d$ übernommen. Offensichtlich befriedigt jede umkehrbare Lösung des Differentialgleichungssystems (2) das gemischte System (2), (7), und auch umgekehrt.

Wir können (2) mit Hilfe von (7b) auf eine Form bringen, welche die Übersichtlichkeit der folgenden Relationen in beträchtlichem Maße fördert. Indem wir den aus (7b) sich ergebenden Wert von $p_c{}^g{}_d$ in (2) einsetzen, erhalten wir

$$A^{gh}{}_{klt}\, p^k_c\, p^l_d\, p^t_f = \bar{A}^{ab}{}_{cdf}\, p^g_a\, p^h_b, \tag{8}$$

wo

$$A^{gh}{}_{klt} \equiv (n+1)\gamma_{kl}{}^{gh}{}_t - \left(\delta^h_l\, G_k{}^g{}_t + \delta^g_k\, G_l{}^h{}_t\right) \tag{9}$$

ist. $\{\mathfrak{P}, \gamma\}$ und $\{\mathfrak{P}, \bar{\gamma}\}$ sind dann und nur dann äquivalent, wenn (7) und (8) ein Lösungssystem besitzt, wo (1) umkehrbar ist.

Der Satz von J. M. THOMAS und O. VEBLEN [3] bezüglich der Lösbarkeit des gemischten Systems von Typ (6) sagt folgendes aus: Wir bilden die Integrabilitätsbedingungen von (6a) und die partiellen Ableitungen nach $\bar{x}^s$ von (6b), wobei wir auch (6a) berücksichtigen. So erhalten wir gewisse Relationen

$$F_1^{\lambda_1}(\Theta, \bar{x}) = 0 \qquad (\lambda_1 = 1, 2, \ldots, S_1).$$

Aus diesen letzteren entsteht nun durch ein ähnliches Verfahren eine neue Gleichungsgruppe. Indem wir dieses Verfahren fortsetzen, gelangen wir zu den Gleichungen

$$F_1^{\lambda_1}(\Theta, \bar{x}) = 0, \ldots, F_N^{\lambda_N}(\Theta, \bar{x}) = 0. \tag{10}$$

Nach dem erwähnten Satz ist für die Lösbarkeit von (6) die Existenz eines solchen N notwendig und hinreichend, daß die N ersten Gleichungsgruppen von (10) für jedes $\bar{x}$ in einem Gebiet ein verträgliches Gleichungssystem bilden, und daß $F_{N+1}^{\lambda_{N+1}} = 0$ schon eine Folge der vorangehenden Gleichungen ist. Bezeichnet p die Anzahl der unabhängigen Gleichungen von (10), und M die Anzahl der Unbekannten des gemischten Systems, so enthalten die Lösungsfunktionen $M - p$ willkürliche Parameter.[2]

Zwecks der Anwendung dieses Satzes bilden wir nun mit Rücksicht auf (7) die Integrabilitätsbedingungen von (7). Die Integrabilitätsbedingung von (7a) ist mit Rücksicht auf (7b)

$$\bar{S}_{f\,c}^{\,a}\, p_a^g = S_{k\,t}^{\,g}\, p_c^k\, p_d^t, \tag{11}$$

wo

$$S_{k\,t}^{\,g} \equiv \frac{1}{n+1}\left(G_{k\,t}^{\,g} - G_{t\,k}^{\,g}\right) \tag{12}$$

ist. Die Integrabilitätsbedingung von (7b) ist, mit Rücksicht auf (7b) und (7a)

$$\bar{R}^{a}{}_{bcd}\, p_a^s = R^s{}_{hji}\, p_b^h\, p_c^j\, p_d^i, \tag{13}$$

wo[3]

$$R^s{}_{hji} \equiv \frac{1}{(n+1)^2}\left(G_h{}^v{}_j\, G_v{}^s{}_i - G_i{}^v{}_j\, G_v{}^s{}_h\right) + \frac{1}{n+1}\left(\frac{\partial G_j{}^s{}_h}{\partial x^i} - \frac{\partial G_j{}^s{}_i}{\partial x^h}\right) \tag{14}$$

ist. Die rechte Seite von (7c) ist die partielle Ableitung von $\frac{\partial p_c^g}{\partial \bar{x}^d}$ nach $\bar{x}^m$. Es drückt die Integrabilitätsbedingung von (7c) die Gleichheit der zweiten partiellen Ableitung von p_c^g hinsichtlich $\bar{x}^d$ und $\bar{x}^m$, mit der in entgegengesetzten

[2] Bezüglich einer solchen Fassung des Satzes siehe [2], S. 16.

[3] Die Zweckmäßigkeit des Einschaltens des Faktors $\frac{1}{n+1}$ geht aus Satz 9 hervor.

Reihenfolge genommenen zweiten partiellen Ableitung aus, die für eine zweimal stetig differenzierbare Funktion eine Identität ist.

Bevor wir (8) nach $\bar{x}^s$ ableiten, bemerken wir, daß auf Grund des Transformationsgesetzes (5), die Größe $\frac{1}{n+1} G_i{}^j{}_k$ zur Bildung einer kovarianten Ableitung geeignet ist.[4] Die partielle Ableitung von (8) nach $\bar{x}^s$ ist mit Rücksicht auf (7b) und (7a)

(15) $$\bar{A}^{ab}{}_{cdf|s}\, p^g_a\, p^h_b = A^{gi}{}_{klt|p}\, p^k_c\, p^l_d\, p^t_f\, p^p_s,$$

wenn wir die mit Hilfe von $\frac{1}{n+1} G$ nach $\bar{x}^s$ gebildete kovariante Ableitung mit einem nach einem Strich gesetzten s bezeichnen. Es ist leicht einzusehen, daß die hinsichtlich $\bar{x}^v$ gebildete partielle Ableitung eines Ausdrucks, der das tensorielle Transformationsgesetz einer beliebigen Größe ausdrückt, mit Rücksicht auf (7b) und (7a) einen solchen Ausdruck ergibt, der das tensorielle Transformationsgesetz der mit Hilfe von $\frac{1}{n+1} G$ gebildeten kovarianten Ableitung dieser Größe beschreibt. Wird (15) anstatt den jetzt $F_1^{\lambda_1} = 0$ entsprechenden Relationen $F_2^{\lambda_2} = 0$ entsprechenden Relationen gereiht (was keine wesentliche Änderung bedeutet), so erhalten wir mit Hilfe des Kriteriums von THOMAS und VEBLEN die folgende Bedingung:

SATZ 5. *Notwendig für die Äquivalenz der Räume $\{\mathfrak{P}, \gamma\}$ und $\{\mathfrak{P}, \bar{\gamma}\}$ ist die Existenz einer Zahl N von der Art, daß die Relationen, die das tensorielle Transformationsgesetz der mit Hilfe von $\frac{1}{n+1} G$ gebildeten ersten $N-1$ kovarianten Ableitungen der Tensoren A, S, R zum Ausdruck bringen, ein verträgliches Gleichungssystem für jedes $\bar{x}$ in einem n-dimensionalen Gebiet bilden, und daß die Relationen, die das tensorielle Transformationsgesetz der N-ten kovarianten Ableitungen ausdrücken, schon Folgerungen der Vorangehenden sind.*

Ist die Lösung des gemischten Systems (7), (8) eine solche, daß man in jedem Punkt $\bar{x}_0$ eines gewissen Bereiches $\mathfrak{B}$ ein Wertesystem $\overset{1}{c_0}, \ldots, \overset{M-p}{c_0}$ derart finden kann, daß die Determinante $|p^i_a(\bar{x}_0, c_0)|$ der Lösungsfunktionen $p^i_a(\bar{x}, \overset{1}{c_0}, \ldots, \overset{M-p}{c_0})$ nicht verschwindet, bilden die in den Lösungen vorkommenden Funktionen $x^i = x^i(\bar{x}, c_0)$ für jedes $\bar{x}$ aus $\mathfrak{B}$ für die entsprechenden c_0 ein umkehrbares Funktionensystem, d. h. die zwei Räume sind äquivalent. Ob ein solcher c_0 existiert oder nicht, ist bei bekannten $p^i_a(\bar{x}, c)$ sehr leicht

[4] Siehe [2], S. 3.

zu entscheiden. Wir wollen aber auch eine solche Bedingung finden, die ohne Kenntnis der Lösungen die Umkehrbarkeit der $x^i = x^i(\bar{x}, c)$ sichert. Eine solche Bedingung ist $p = 0$, wo p (nach den Vorangehenden) die Anzahl der unabhängigen Relationen ist, die jetzt den Gleichungen (10) entsprechen. Dann enthalten nämlich die M Lösungsfunktionen M beliebige Parametern, und diese sind in einem beliebig festgehaltenen Punkt $\bar{x}_0$ so wählbar, daß die Lösungsfunktionen, und so auch die $p^i_a(\bar{x}, c)$ für diese Parameterwerte beliebig vorgegebene Funktionenwerte annehmen. Es gilt daher der

SATZ 6. *Für die Äquivalenz der Räume* $\{\mathfrak{P}, \gamma\}$ *und* $\{\mathfrak{P}, \bar{\gamma}\}$ *ist es hinreichend, wenn außer der Lösbarkeit des gemischten Systems* (7), (8), *noch die in der Lösung auftretenden Konstanten in jedem Punkt* $\bar{x}_0$ *eines gewissen Bereiches so wählbar sind, daß* $|p^i_a(\bar{x}_0, c_0)| \neq 0$ *ist. Dies ist sicher der Fall, wenn* $p = 0$ *ist.*

§ 4. Krümmung und Torsion

1. DEFINITION. Wir nennen den Raum $\{\mathfrak{P}, \gamma\}$ affin, wenn es ein solches Koordinatensystem gibt, in dem sämtliche Koeffizienten der γ überall verschwinden.

Gemäß dem vorangehenden Paragraphen besteht die Bedingung dafür, daß der $\{\mathfrak{P}, \gamma\}$ ein affiner Raum im Sinne der hier gegebenen Definition ist, darin, daß das Differentialgleichungssystem

$$p_c{}^g{}_{/f}\, p_a^h + p_a{}^h{}_{/f}\, p_c^g + \gamma_{kl}{}^{gh}{}_t(x)\, p_c^k\, p_a^l\, p_f^t = 0 \tag{1}$$

eine umkehrbare Lösung bezüglich der Funktionen

$$x^i = x^i(\bar{x}) \tag{2}$$

besitzt.

SATZ 7. *Eine notwendige und hinreichende Bedingung dafür, daß* $\{\mathfrak{P}, \gamma\}$ *einen affinen Raum bildet, ist das Verschwinden der Krümmungstensoren* R *und* A, *und des Torsionstensors* S.

R, A und S sind wirklich Tensoren. Sind nämlich die $x^i = x^i(\bar{x})$ nicht unbekannte Funktionen, wie das im vorigen Paragraphen der Fall war, sondern Funktionen, die eine Transformation beschreiben, so sind (§ 3, (13)), (§ 3, (8)) und (§ 3, (11)) nicht Bedingungsgleichungen, sondern Transformationsgesetze. Dies bedeuten aber eben, daß R, A und S Tensoren sind.

Das Verschwinden der erwähnten Tensoren ist wirklich notwendig. Die Komponenten von γ, oder deren Ableitungen nach (§ 3, (14)), (§ 3, (9)), (§ 3, (12)) und (§ 3, (4)) kommen nämlich in jedem Glied der Ausdrücke für R, A und

S vor. Ist $\{\mathfrak{P}, \gamma\}$ affin, so gibt es nach unserer Definition ein Koordinatensystem, in welchem sämtliche Koeffizienten der γ überall verschwinden. Dann verschwinden aber hier auch die Ableitungen der γ, und so auch die Tensoren R, A, S. Dies ist aber dann in jedem Koordinatensystem der Fall, da sie Tensoren sind.

Anknüpfend an die beiden Sätze des vorigen Paragraphen zeigen wir jetzt, daß im Falle des Verschwindens der R, A und S $\{\mathfrak{P}, \gamma\}$ äquivalent mit jenem, auf das Koordinatensystem $\bar{x}$ bezogenen $\{\mathfrak{P}, \bar{\gamma}\}$ ist, für welchen $\bar{\gamma}$ identisch verschwindet. Dies wird im Sinne unserer Definition den hinreichenden Teil unseres Satzes beweisen.

Das jetzt zu lösende gemischte System reduziert sich auf das Differentialgleichungssystem

$$(3)\qquad \begin{aligned} &\text{a)} \quad \frac{\partial x^i}{\partial \bar{x}^a} = p^i_a(x), \\ &\text{b)} \quad \frac{\partial p^g_c}{\partial \bar{x}^d} = -\frac{1}{n+1} G_k{}^g{}_t(x)\, p^k_c p^t_d. \end{aligned}$$

Das in (§ 3, (7b)) vorkommende $\bar{G}$ fällt wegen des Verschwindens der $\bar{\gamma}$ weg. Die linke Seite von (§ 3, (8)) verschwindet wegen unserer Bedingung, und für die rechte Seite gilt dies wegen des Verschwindens von $\bar{\gamma}$. (§ 3, (7c)) ist zu vernachlässigen, wie wir darauf schon im vorigen Paragraphen hingewiesen haben. In diesem Falle existiert wegen des vorausgesetzten Verschwindens von R und S eine Zahl N, welche die im Satz 5 geforderte Eigenschaft besitzt, und zwar ist $N = 1$. — Da die dem (§ 3, (10)) jetzt entsprechenden Gleichungen lauter Identitäten sind, so ist die Anzahl der unabhängigen Gleichungen gleich Null. So geht auch die Bedingung des Satzes 6 in Erfüllung, was schon die Äquivalenz von $\{\mathfrak{P}, \gamma\}$ und $\{\mathfrak{P}, \bar{\gamma}\}$ sichert. Damit haben wir auch den hinreichenden Teil des Satzes 7 bewiesen.

2. Nunmehr möchten wir auf die geometrische Bedeutung des Tensors A hinweisen.

SATZ 8. *Die hinreichende und notwendige Bedingung dafür, daß der Raum $\{\mathfrak{P}, \gamma\}$ sich auf einen $\{\mathfrak{P}, \Gamma\}$ reduziert, ist das Verschwinden des Krümmungstensors A.*

Das Verschwinden von A ist hinreichend. Dann ist nämlich (§ 2, (4)) lösbar, und dies ist gemäß dem Satz 2 zur Reduktion vom $\{\mathfrak{P}, \gamma\}$ auf einen $\{\mathfrak{P}, \Gamma\}$ schon hinreichend. Die Lösung ist $\Gamma_j{}^i{}_k = \frac{1}{n+1} G_j{}^i{}_k$. — Das Verschwinden von A ist auch notwendig. Reduziert sich nämlich $\{\mathfrak{P}, \gamma\}$ auf einen

$\{\mathfrak{P}, \Gamma\}$, d. h. ist (§ 2, (4)) lösbar, so gilt — wenn man diese Werte der γ in (§ 3, (4)) einsetzt —

$$G_{k\ t}^{\ h} = (n+1)\,\Gamma_{k\ t}^{\ h}. \tag{4}$$

Indem wir jetzt die nach (§ 2, (4)) genommenen Werte der γ und die nach (4) genommenen Werte der G aber in den Ausdruck (§ 3, (9)) für A einsetzen, erhalten wir Null.

Aus dem Gesagten, aus (4), sowie aus den Relationen (§ 3, (14)) und (§ 3, (12)), die sich aus der Definition von R und von S ergeben, folgt der

SATZ 9. *R und S reduzieren sich im Falle des Verschwindens des Tensors A auf die Krümmung und auf die Torsion eines durch das Γ erzeugten (gewöhnlichen) affinzusammenhängenden Raumes.*

Ist $\{\mathfrak{P}, \gamma\}$ affin, so verschwindet nach Satz 7 der Tensor A. Dann reduziert sich $\{\mathfrak{P}, \gamma\}$ nach Satz 8 auf einen Raum $\{\mathfrak{P}, \Gamma\}$, welcher nach Satz 4 der einzige derartige Raum ist. $\{\mathfrak{P}, \Gamma\}$ ist aber ein Teil des über der Menge $\mathfrak{T}$ sämtlicher Tensoren durch das Γ erzeugten gewöhnlichen affinzusammenhängenden Raumes (welches wir mit $\{\mathfrak{T}, \Gamma\}$ bezeichnen). Wegen des Verschwindens der Tensoren R und S (Satz 7) verschwindet aber nach Satz 9 sowohl die Krümmung als auch die Torsion dieses $\{\mathfrak{T}, \Gamma\}$, also ist dies ein affiner Raum in gewöhnlichem Sinne. Daher gilt der

SATZ 10. *Ist $\{\mathfrak{P}, \gamma\}$ in dem am Anfang dieses Paragraphen gegebenen Sinne affin, so ist er ein (die Tensoren gerader Stufe enthaltender) Teil eines gewöhnlichen affinzusammenhängenden Raumes.*

3. Endlich untersuchen wir die Bedingungen der Existenz eines absolut parallelen Tensorfeldes.

Diese Untersuchung verläuft ganz analog zur Untersuchung der entsprechenden Frage für den Raum $\{\mathfrak{T}, \Gamma\}$, und hängt mit der Integrabilität des Differentialgleichungssystems

$$\frac{\partial T^{ij}(x)}{\partial x^t} + \gamma_{kl\ \ t}^{\ \ ij}(x)\, T^{kl}(x) = 0$$

zusammen, mit dem Unterschied, daß hier anstatt des dem Raume $\{\mathfrak{T}, \Gamma\}$ entsprechenden Krümmungstensor R, der Tensor

$$B^{ij}_{\ \ klvt} \equiv \frac{\partial \gamma_{kl\ \ v}^{\ \ ij}}{\partial x^t} - \frac{\partial \gamma_{kl\ \ t}^{\ \ ij}}{\partial x^v} + \gamma_{rm\ \ t}^{\ \ ij}\,\gamma_{kl\ \ v}^{\ \ rm} - \gamma_{rm\ \ v}^{\ \ ij}\,\gamma_{kl\ \ t}^{\ \ rm}$$

auftritt, dessen Verschwinden notwendig und hinreichend dafür ist, daß wir jede Komponente des absolut parallelen Tensorfeldes T^{ij} oder T_{kl} in einem

Punkt x_0 frei wählen können. Die Relation geht im Falle der Reduktion des $\{\mathfrak{P}, \gamma\}$ auf einen $\{\mathfrak{P}, \Gamma\}$ in

$$\delta^i_k R^j_{lst} + \delta^j_l R^i_{kst} = 0$$

über, was auch ohne das Verschwinden des R in Erfüllung gehen kann.

§ 5. Geometrischer Hintergrund des Falles der Reduktion

Wir haben im § 2 und im Satz 8 mehrere verschiedene notwendige und hinreichende analytische Bedingungen dafür gegeben, daß ein $\{\mathfrak{P}, \gamma\}$ Raum sich auf einen $\{\mathfrak{P}, \Gamma\}$ reduziere. Jetzt geben wir endlich eine rein geometrische Bedingung für diese Reduktion.

Wir haben den affinen Zusammenhang des Raumes $\{\mathfrak{P}, \gamma\}$, welcher allgemeiner ist als der für $\{\mathfrak{P}, \Gamma\}$, durch die unmittelbare affine Zuordnung der Tensoren T^{ij} erreicht. Diese Tensoren bilden in einem Punkt unseres Raumes ein Vektorfeld von n^2 Dimension, genauer einen mit sich selbst gebildeten Produktraum des Tangentialraumes E_n.[5] Sind ξ und η Elemente von E_n, so nennen wir das Element des Produktraumes vom Typ

$$V^{ij} = \xi^i \eta^j \tag{1}$$

dekomponierbar. Offensichtlich sind nur einige spezielle Elemente des Produktraumes (nur einige spezielle Tensoren zweiter Stufe) dekomponierbar.

Dem Element $T^{ij}(x)$ des Punktes x ist mit Hilfe des Operators ϑ dasjenige Element des Punktes $x+dx$ zugeordnet, dessen Komponenten bis auf in dx lineare Glieder die

$$T^{ij}(x) - \gamma_{kl}{}^{ij}{}_t(x)\, T^{kl}(x)\, dx^t$$

sind. Es kann vorkommen, daß diese Zuordnung die dekomponierbaren Elemente der benachbarten Produkträume $T^{ij}(x)$ und $T^{ij}(x+dx)$ einander zuordnet. Wir zeigen den

[5] Sind die Basisvektoren des n-dimensionalen Tangentialraumes E_n die $\vec{e}_i$ $(i=1, 2, \dots, n)$, so ist der n^2-dimensionale Produktraum definitionsgemäß durch die Paare $\vec{e}_i \times \vec{e}_j$ aufgespannt, und er unterscheidet sich von einem n^2-dimensionalen Vektorraum darin, daß hier nicht alle Basistransformationen $\vec{e}_{i'} \times \vec{e}_{j'} = A^{i\,j}_{i'j'}(\vec{e}_i \times \vec{e}_j)$ $(|A^{i\,j}_{i'j'}| \neq 0)$ erlaubt sind, sondern nur solche, wo $A^{i\,j}_{i'j'}$ die Form $A^i_{i'} A^j_{j'}$ $(|A^i_{i'}| \neq 0)$ hat, und eine solche Transformation muß dann in unserem Produktraum durchgeführt werden, wenn der Raum E_n die Transformation $\vec{e}_{i'} = A^i_{i'} \vec{e}_i$ erleidet.

Für den Begriff des Produktraumes siehe das Büchlein Lichnerowicz, *Éléments du calcul tensoriel* (Paris, 1950).

SATZ 11. *Eine notwendige und hinreichende Bedingung dafür, daß der Raum $\{\mathfrak{P}, \gamma\}$ sich auf den Raum $\{\mathfrak{P}, \Gamma\}$ reduziere, ist das, daß der Operator ϑ jedem dekomponierbaren Element zweiter Ordnung der Menge $\mathfrak{P}$ ein Element des benachbarten Punktes zuordnet, das in dem Sinne dekomponierbar ist, daß dieses Element bis auf in dx lineare Glieder das Produkt zweier Vektoren ist.*

Dieser Satz charakterisiert den Fall der Reduktion durch eine unmittelbare geometrische Eigenschaft.

Die gegebene Bedingung ist notwendig. Tatsächlich ist im Falle, wo ϑ sich auf D reduziert, der

$$V^{ij}(x)-[\Gamma_{k}{}^{i}{}_{t}(x)\, V^{kj}(x)+\Gamma_{k}{}^{j}{}_{t}(x)\, V^{ik}(x)]\, dx^{t}, \tag{2}$$

der dem nach (1) dekomponierbaren $V^{ij}(x)$ zugeordnet ist, im Sinne des Satzes wieder dekomponierbar. (2) ist nämlich unter Vernachlässigung von Gliedern, die in dx von zweiter Ordnung sind, ein tensorielles Produkt der Vektoren, die im Punkt $x+dx$ die Komponenten

(3) a) $\xi^{i}(x)-\Gamma_{k}{}^{i}{}_{t}(x)\,\xi^{k}(x)\,dx^{t}$ und b) $\eta^{j}(x)-\Gamma_{k}{}^{j}{}_{t}(x)\,\eta^{k}(x)\,dx^{t}$

besitzen, und dasselbe ist auch im Falle von dekomponierbaren kovarianten Tensoren zweiter Ordnung der Fall.

Zwecks Vorbereitung des zweiten Teiles des Beweises zeigen wir, daß für einen in der Form (1) dekomponierbaren Tensor V^{ij}, der kein Nulltensor ist, die in allen anderen Dekompositionen

$$V^{ij}=\sigma^{i}\varrho^{j} \tag{4}$$

vorkommenden σ^{i} und ϱ^{j} die Form

$$\sigma^{i}=\lambda\xi^{i}, \qquad \varrho^{j}=\frac{1}{\lambda}\eta^{j} \qquad (\lambda\neq 0) \tag{5}$$

besitzen. Es sei V^{ij} ein derartiger Tensor, und es sei $V^{i_0j_0}$ irgendeine nicht verschwindende Komponente von V^{ij}. Es ist dann $V^{i_0j_0}=\xi^{i_0}\eta^{j_0}=\sigma^{i_0}\varrho^{j_0}\neq 0$ und $\xi^{i_0}, \eta^{j_0}, \sigma^{i_0}, \varrho^{j_0}\neq 0$. Daher ist $\sigma^{i_0}=\lambda\xi^{i_0}$ nur für den einzigen $\lambda=\dfrac{\sigma^{i_0}}{\xi^{i_0}}\neq 0$ richtig, und so ist notwendigerweise $\varrho^{j_0}=\dfrac{1}{\lambda}\eta^{j_0}$. Wir nehmen an, daß es solche $V^{i_1j_1}\neq 0$ gibt, wo σ^{i_1} und ϱ^{j_1} nicht in der Form $\sigma^{i_1}=\lambda\xi^{i_1}$, $\varrho^{j_1}=\dfrac{1}{\lambda}\eta^{j_1}$ darstellbar sind, sondern $\sigma^{i_1}=\mu\xi^{i_1}$, $\varrho^{j_1}=\dfrac{1}{\mu}\eta^{j_1}$ $(\mu\neq 0)$ ist. Wir betrachten die Komponente $V^{i_0j_1}=\xi^{i_0}\eta^{j_1}=\sigma^{i_0}\varrho^{j_1}$, welche wegen der für die σ^{i_0} und ϱ^{j_1} bestehenden Relationen auch $\lambda\xi^{i_0}\dfrac{1}{\mu}\eta^{j_1}$ gleich sein muß. Dies ist aber nur im

Falle $\lambda=\mu$ möglich. So besteht für die nichtverschwindenden Komponenten von V (5). — Ist $V^{i_1 j_1}=\xi^{i_1}\eta^{j_1}=\sigma^{i_1}\varrho^{j_1}=0$, so ist in den zwei mittleren Ausdrücken mindestens je ein Faktor gleich Null. Für die einander entsprechenden verschwindenden Faktoren (einander entsprechend sind ξ^{i_1} und σ^{i_1} bzw. η^{j_1} und ϱ^{j_1}) besteht die entsprechende Relation von (5). Für einander entsprechende nichtverschwindende Faktoren ist das Bestehen der entsprechenden Relation von (5) ähnlich wie soeben einzusehen. Es würde noch der Fall übrigbleiben, wo ein verschwindender und ein nichtverschwindender Faktor einander entsprechen. Es sei z. B. $\xi^{i_1}\neq 0$ und $\sigma^{i_1}=0$. Dann ist $\eta^{j_1}=0$. Dies führt aber zu einem Widerspruch, weil dann wegen $\sigma^{i_1}=0$ jedes Element der i_1-ten Zeile der $\|V^{ij}\|$ verschwindet, was wegen $\xi^{i_1}\neq 0$ nur so möglich ist, wenn $\eta^j=0$ für $j=1,2,\ldots,n$ ist. Dann sind aber alle V^{ij} gleich Null, in Widerspruch zu unserer Bedingung. So ist (5) für jede Dekomposition eines nicht identisch verschwindenden Tensors gültig.

Wir zeigen, daß die Bedingung des Satzes auch hinreichend ist. Es sei also $V^{ij}(x)$ in der Form (1) dekomponierbar, und es sei der ihm im Punkte $x+dx$ zugeordnete Tensor

$$V^{ij}(x)-\gamma_{kl}{}^{ij}{}_t(x)\,V^{kl}(x)\,dx^t \tag{6}$$

im Sinne des Satzes wieder dekomponierbar, d. h. er sei unter Berücksichtigung der bis auf in dx linearen Glieder als Produkt der in $x+dx$ genommenen Komponenten zweier Vektorfelder ξ^* und η^* darstellbar. Das von dx freie Glied von (6) kann nur als das Produkt der von dx freien Glieder von ξ^* und η^* entstehen. Aber jede Dekomposition von V^{ij} hat die Form (5). So muß das von dx freie Glied der zwei Vektoren $\lambda\xi^i$ bzw. $\frac{1}{\lambda}\eta^j$ sein. Der in dx lineare Teil von (6) kann nur durch Multiplikation des von dx freien Gliedes des einen der ξ^* und η^* mit dem in dx linearen Glied des anderen entstehen. Um bei einem solchen Produkt einen Ausdruck vom gleichen Typ als das zweite Glied von (6) zu erhalten, also einen solchen, in dem außer dx und den entsprechenden Koeffizienten nur die Komponenten des V vorkommen, müssen die ξ^{*i} bzw. η^{*j} im $x+dx$ mit λ bzw. mit $\frac{1}{\lambda}$ multiplizierten Werte von (3a) bzw. (3b) haben, mit irgendwelchen Koeffizienten $\Gamma_p{}^r{}_s$. Das Produkt derselben gibt unter Berücksichtigung der bis auf in dx linearen Glieder den Ausdruck (2). Daher muß

$$\gamma_{kl}{}^{ij}{}_t\,V^{kl}dx^t=(\Gamma_k{}^i{}_t\,V^{kj}+\Gamma_l{}^j{}_t\,V^{il})\,dx^t$$

sein, d. h. es muß

$$(\gamma_{kl}{}^{ij}{}_t-\delta_l^j\,\Gamma_k{}^i{}_t-\delta_k^i\,\Gamma_l{}^j{}_t)\,V^{kl}\,dx^t=0$$

für jede dx, und für jeden dekomponierbaren V^{kl} gelten. Dies bedeutet abe das Verschwinden des in Klammer stehenden Ausdruckes. Die γ müsse daher die Relation (§ 2, (4)) erfüllen. Dann reduziert sich aber nach Satz der Operator ϑ auf D, d. h. $\{\mathfrak{P}, \gamma\}$ auf einen $\{\mathfrak{P}, \Gamma\}$.

Endlich spreche ich Herrn Professor O. VARGA für seine Unterstützun bei meiner Arbeit und besonders für eine auf das Ergebnis des § 5 führend Anleitung meinen besten Dank aus.

(Eingegangen am 21. Juli 1959.)

Literaturverzeichnis

[1] E. BOMPIANI, Le conessioni tensoriali, *Atti Acad. dei Lincei*, **1** (1946), S. 478—482.
[2] L. P. EISENHART, *Non-Riemannian geometry* (New York, 1927).
[3] J. M. THOMAS and O. VEBLEN, Projektive invariants of affine geometry of paths, *Annal of Math.*, **27** (1926), S. 279—296.
[4] O. VARGA, Über affinzusammenhängende Mannigfaltigkeiten von Linienelementen, ins besondere deren Äquivalenz, *Publ. Math. Debrecen*, **1** (1949), S. 7—17.
[5] H. WEYL, Reine Infinitesimalgeometrie, *Math. Zeitschrift*, **2** (1918), S. 384—411.

A RULE OF DUALISM IN MATHEMATICAL STATISTICS

By

K. SARKADI (Budapest)

(*Presented by* A. RÉNYI)

Introduction

The present paper is related to the papers of H. STEINHAUS and J. ODERFELD ([18], [9], [10]). These authors express the opinion that Bayes' method in statistical inference has to be rehabilitated. In these communications a prominent role is played by the "rule of dualism". This rule, discussed by ODERFELD in detail ([9], [10]), states a connection between the confidence limits and the inverse probability limits of the fraction defective in case the a priori distribution is uniform or of Beta type.

The present paper gives generalizations of the rule of dualism (Sections 1—5); it contains, further, some remarks on the statistical method of an a priori Beta distribution proposed by ODERFELD (Section 6).

Most of the contents of this paper has been already published in Hungarian ([13], [14]).

In case of an a priori uniform distribution the rule of dualism is based on the following equality (already known, see e. g. [3]):

$$\mathbf{P}\{\zeta < p \mid \varkappa_n = k\} = \mathbf{P}\{\varkappa_{n+1} \geqq k+1 \mid \zeta = p\}. \tag{1}$$

Here ζ denotes the fraction defective, $\varkappa_n$ the number of defective items found in a sample of size n (i. e. $\varkappa_n$ is a variate of binomial distribution with parameters n and ζ); k is a positive integer; $0 < p < 1$.

The first generalization of (1) concerns the case of a hypergeometric sample distribution which represents the real situation when examining finite lots. In addition, we shall show that it is possible to establish a rule of dualism even in case of more than two alternatives and in case of continuous production processes when the number of defects in a lot obeys a Poisson distribution.[1]

In contrary to ODERFELD, our proof is based on general probabilistic arguments instead of direct computations.

Section 5 deals with the problem of applying the rule of dualism in case of a non-uniform a priori distribution. ODERFELD showed [9] that, in

[1] These problems were suggested by A. RÉNYI.

the case of an a priori Beta distribution, a modification of the rule of dualism holds. We show that also in case of a mixture of Beta distributions a form of the rule of dualism is valid; although more complicated than the former one. If the number of the components is not large, then the expressions may have practical application. This is of great importance, since — as it is well known — any probability distribution with continuous density function defined on the interval (0, 1) can be approximated with any prescribed accuracy by a mixture of a finite number of Beta distributions.

In [9] the method of a priori Beta distribution is suggested for the case when many lots are inspected one after the other and, after inspecting sufficiently large number of lots, the empirical data justify the assumption of an a priori Beta distribution; the two parameters can be determined on the basis of the data. As it is pointed out (Section 6), ODERFELD does not take into account that the empirical data consist of sample results; further it is shown that sampling errors can produce absurd results relating to the moments. These difficulties can be avoided through the modifications proposed here.

1. The hypergeometric distribution

If the finite lot contains N elements, then, as proved below, instead of (1) the more precise equality holds:

THEOREM 1.

$$\mathbf{P}\left\{\zeta \leqq \frac{M}{N}\,\middle|\, \varkappa_{nN}=k\right\} = \mathbf{P}\left\{\varkappa_{n+1,\,N+1} \geqq k+1 \,\middle|\, \zeta=\frac{M+1}{N+1}\right\}. \tag{2}$$

Here the second index of $\varkappa$ denotes the size of the lot from which the sample was taken. M is an integer, the other notations are according to former use. In this case the a priori assumption is as follows: The fraction defective takes on the values k/N ($k=0, 1, \ldots, N$) each with probability $1/(N+1)$. This form of the equality is given in the quality control handbook of L. E. SIMON [16], with reference to E. C. MOLINA [8]. On the cited page, however, neither the theorem nor the proof can be found. The problem is dealt with by JORDAN [7] and ROMANOVSKY [12], too, the latter paper is known to the author through the review only.

Our proofs will be carried out with the aid of models:

In order to prove (2), let us imagine the following model: the elements of a lot of size $N+1$ are labelled from 1 to $N+1$. From this lot a sample of size $n+1$ is taken. Let $\xi_1, \xi_2, \ldots, \xi_{n+1}$ denote the original serial numbers of the elements. We note that in this section "sample" always denotes a sample without replacement. Consequently, the set does not contain identical ξ's.

Let us rearrange the ξ's in ascending order of magnitude and denote by $\xi_{(1)}, \xi_{(2)}, \ldots, \xi_{(n+1)}$ the ordered values.

It is easy to see that the following equality holds:

$$\mathbf{P}\{\xi_{(k+1)} \leqq M+1 \mid \xi_{(k+1)} = \xi_1\} = \mathbf{P}\{\xi_{(k+1)} \leqq M+1\}. \tag{3}$$

$\xi_{(k+1)}$ is namely a symmetrical function of the equivalent random variables $\xi_1, \xi_2, \ldots, \xi_{n+1}$ and thus, putting another ξ_i instead of ξ_1, the probability on the left-hand side of (2) does not change, consequently it agrees with the unconditional probability.

In addition, it can be shown that

$$\mathbf{P}\{\xi_{(k+1)} \leqq M+1\} = \mathbf{P}\left\{\varkappa_{n+1, N+1} \geqq k+1 \mid \zeta = \frac{M+1}{N+1}\right\}. \tag{4}$$

As a matter of fact, let us denote the first $M+1$ elements of the lot as "defective" and the remaining $N-M$ as "good" ones. The fact, that under such conditions in the sample of size $n+1$ taken from the lot there are at least $k+1$ defective ones, is the same as that the $(k+1)$-st ordered element of the sample has a serial number at most $M+1$, i. e. it is still a defective one. The two sides of (4) give the probabilities of these two equivalent events.

Finally, we shall prove that

$$\mathbf{P}\{\xi_1 \leqq M+1 \mid \xi_{(k+1)} = \xi_1\} = \mathbf{P}\left\{\zeta \leqq \frac{M}{N} \,\middle|\, \varkappa_N = k\right\}. \tag{5}$$

For this purpose let us consider again a labelled lot of size $N+1$ and draw from it the sample element ξ_1 at random. Let us call the elements with serial numbers smaller than ξ_1 in the remaining lot of size N "defectives", the other elements "good". Thus we have obtained a lot with an a priori uniformly distributed fraction defective, according to our above assumption. Now we take a sample $\xi_2, \xi_3, \ldots, \xi_{n+1}$ of size n from the lot of size N. The event, that in the sample $\xi_1, \xi_2, \ldots, \xi_{n+1}$ of size $n+1$ ξ_1 will be the $(k+1)$-st element according to the order of magnitude, is equivalent to the event that in the sample $\xi_2, \xi_3, \ldots, \xi_{n+1}$ of size n there will be exactly k defective pieces, i. e. the event $\xi_{(k+1)} = \xi_1$ agrees with the event $\varkappa_{nN} = k$. In addition, since the fraction defective of the lot is $(\xi_1 - 1)/N$, the event $\xi_1 \leqq M+1$ means that the fraction defective of the lot of size N is $\zeta \leqq M/N$. From equations (3), (4) and (5) Theorem 1 follows. So we have proved the rule of dualism also in the case of a finite lot. Using a limiting process, (1) may be obtained from (2), too.

2. The polyhypergeometric distribution

It can be proved that if there are instead of a simple alternative $s+1$ different classes, the rule of dualism may be extended in the following form:

THEOREM 2.

$$\mathbf{P}\left\{\prod_{r=1}^{s}\left(\sum_{i=1}^{r}\zeta_i \leq \sum_{i=1}^{r}\frac{M_i}{N}\right)\Bigg|\prod_{r=1}^{s}(\varkappa_{rnN}=k_r)\right\}=$$

$$(6)\qquad =\mathbf{P}\left\{\prod_{r=1}^{s}\left(\sum_{i=1}^{r}\varkappa_{i,\,n+s,\,N+s} \geqq \sum_{i=1}^{r}k_i+r\right)\Bigg|\prod_{r=1}^{s}\left(\zeta_r=\frac{M_r+1}{N+s}\right)\right\}.$$

Here ζ_i denotes the proportion of the elements of the i-th class in the lot. $\varkappa_{inN}$ denotes the number of elements belonging to the i-th class in a sample of size n, taken from a lot of size N. The product sign signifies that the event following it holds simultaneously for every value of the index.

Theorem 2 holds in the case of an a priori Bose—Einstein distribution when any distribution of the lot among the $s+1$ classes has the same probability (cf. also [7]).

Here the way of proving is the same as in the preceding section, with the difference that instead of a lot of size $N+1$ a lot of size $N+s$ is chosen as a model. The first s elements of a sample of size $n+s$ taken from above lot will determine the limits of the classes; the remaining elements will give the sample of size n, taken from the lot of size N. The detailed proof may be found in the Hungarian version [13] of the present paper.

3. The polynomial distribution

If N tends to infinity, we obtain from Theorem 2

THEOREM 3.

$$\mathbf{P}\left\{\prod_{r=1}^{s}\left(\sum_{i=1}^{r}\zeta_i \leq \sum_{i=1}^{r}p_i\right)\Bigg|\prod_{r=1}^{s}(\varkappa_{rn}=k_r)\right\}=$$

$$(7)\qquad =\mathbf{P}\left\{\prod_{r=1}^{s}\left(\sum_{i=1}^{r}\varkappa_{i,\,n+s} \geqq \sum_{i=1}^{r}k_i+r\right)\Bigg|\prod_{r=1}^{s}(\zeta_r=p_r)\right\}.$$

Here $\varkappa_{rn}$ denotes the number of elements of the r-th class found in a sample of size n. Theorem 3 holds only in the case of a continuous a priori distribution of the lot, corresponding to the Bose—Einstein distribution. This distribution can be generated by a "random" division of the interval $(0, 1)$ into $s+1$ parts, by means of s mutually independent, uniformly distributed random variables.

Theorem 3 can be proved directly, too, in the same way as Theorem 2, with the difference that the variates have to be defined as mutually independent random variables with uniform distribution in the interval (0, 1).

4. Poisson distribution

The Poisson distribution being the limiting form of the hypergeometric and binomial distributions, the rule of dualism holds in the case, too.

Some difficulties arise, however, already at the formulation of the rule, as according to the previous cases, the parameter of the Poisson distribution must be a priori uniformly distributed on the half line $(0, \infty)$. Such a distribution does not exist in the Kolmogorov theory of probability. A precise meaning for the a posteriori distribution under the assumption of such an a priori distribution may be given, however, as the limiting case of the a posteriori distribution according to the hypothesis of a uniformly distributed parameter on the interval $(0, L)$ when L tends to infinity. This interpretation gives the possibility of generalizing equation (1) in the case of a Poisson distribution:

THEOREM 4.

$$\mathbf{P}\{\lambda < t \mid \varkappa = k\} = \mathbf{P}\{\varkappa \geqq k+1 \mid \lambda = t\}, \tag{8}$$

where $\varkappa$ denotes our random variate following the Poisson distribution, λ its expected value having a priori uniform distribution over the half-line $(0, \infty)$.

Equation (8) can be proved in several ways (see [19], [13]).

On the other hand, the theory of the conditional probability fields due to A. RÉNYI [11] gives the possibility of a direct interpretation of the a priori uniform distribution on the half-line $(0, \infty)$. As shown below, Theorem 4 holds in this sense, too.

In what follows we apply a model of RÉNYI ([11], p. 317).

Let S denote the band $0 \leqq x < +\infty$, $0 \leqq y < 1$ of the plane (x, y), $\mathcal{A}$ the set of all measurable subsets of S, $\mathcal{B}$ the set of all Borel-measurable subsets of S with positive and finite plane measure, and of all sets of positive linear measure lying on some line $x = t$ $(0 \leqq t < +\infty)$. Let us put

$$\mathbf{P}\{A \mid B\} = \begin{cases} \dfrac{m_2(AB)}{m_2(B)} & \text{if } m_2(B) > 0, \\[2ex] \dfrac{m_1(AB)}{m_1(B)} & \text{if } B \text{ is lying on a line } x = t, \end{cases}$$

where $m_2(X)$ denotes the plane Lebesgue measure of X and $m_1(X)$ denotes the linear Lebesgue measure of X.

So we obtain the conditional probability field $[S, \mathcal{A}, \mathcal{B}, \mathbf{P}(A|B)]$ in which we define the variates $\varkappa$ and λ as follows:

Let

$$\lambda = \lambda(x, y) = x,$$

further let $\varkappa = \varkappa(x, y) = 0$ for

$$0 \leqq y < e^{-x} \tag{9}$$

and let $\varkappa = \varkappa(x, y) = k$ for

$$\sum_{j=0}^{k-1} \frac{x^j e^{-x}}{j!} \leqq y < \sum_{j=0}^{k} \frac{x^j e^{-x}}{j!}. \tag{10}$$

Obviously, λ has a regular and uniform probability distribution on the positive real axis and $\varkappa$ has a conditional distribution of Poisson type with expectation λ under the condition of any fixed value of λ:

$$\mathbf{P}\{\varkappa = k | \lambda = l\} = \frac{m_1(\varkappa = k, \lambda = l)}{m_1(\lambda = l)} = \frac{l^k e^{-l}}{k!}. \tag{11}$$

On the other hand, we have

$$\mathbf{P}\{\lambda = l | \varkappa = k\} = \frac{m_2(\lambda < l, \varkappa = k)}{m_2(\varkappa = k)}.$$

The value of the denominator is determined by the measure of the set characterized by the inequality (9) or (10). The value of the numerator is given by the area of the plane characterized by the inequality (9) or (10) and by the inequality $x < l$.

In this way we obtain

$$\mathbf{P}\{\lambda < l | \varkappa = k\} = \frac{\int_0^l \frac{x^k e^{-x}}{k!} dx}{\int_0^\infty \frac{x^k e^{-x}}{k!} dx} = \frac{\int_0^l x^k e^{-x} dx}{k!} \tag{12}$$

On the other hand, we obtain from (11)

$$\mathbf{P}\{\varkappa \geqq k+1 | \lambda = l\} = \sum_{j=k+1}^{\infty} \frac{l^j e^{-l}}{j!}. \tag{13}$$

As the identity of the left sides in (11) and (13) is well known, Theorem 4 is proved.

5. Non-uniform a priori distribution

In practice the a priori distribution of the fraction defective is often approximately known from data regarding lots from the same consignment. In these cases it is suitable for practical reasons to approach the a priori distribution with distributions leading to simple computations. SKELLAM [17] and ODERFELD [9] have shown that if the a priori distribution is of Beta type, i. e. it has a frequency function

$$f(x)=\frac{x^{r}(1-x)^{s-r}}{B(r+1,s-r+1)}, \tag{14}$$

then the rule of dualism holds in the following form:

$$\mathbf{P}\{\zeta<p\,|\,\varkappa_n=k\}=\mathbf{P}\{\varkappa_{n+s+1}\geqq k+r+1\,|\,\zeta=p\}. \tag{15}$$

(Here and in what follows the distribution of the fraction defective of the sample will be regarded to be a binomial one.)

We remark that in case of r and s being integers the above result may be deduced also from the distribution arising from the division of an interval, mentioned in Section 3. Namely, if elements belonging to the first r classes are regarded as defective ones and the others as good ones, the fraction defective will follow a Beta distribution (according to the fact that the r-th value of a sample taken from a uniform distribution has a Beta distribution.)

Consequently, the modification of the proof in Section 3 by omitting the product sign and giving a fixed value to r, results in (15). In the next section we shall return to the practical application.

In what follows we shall give another type of a priori distributions for which a special form of the rule of dualism holds.

THEOREM 5. *Let the a priori distribution be a mixture of Beta distributions, i. e. let its frequency function be*

$$f(x)=\sum_{i=1}^{\infty}c_i\frac{x^{r_i}(1-x)^{s_i-r_i}}{B(r_i+1,s_i-r_i+1)}.$$

In this case we obtain

$$\mathbf{P}\{\zeta<p\,|\,\varkappa_n=k\}=\frac{\displaystyle\sum_{i=1}^{\infty}\frac{c_i(s_i+1)\binom{s_i}{r_i}}{(s_i+n+1)\binom{s_i+n}{r_i+k}}\mathbf{P}\{\varkappa_{n+s_i+1}\geqq k+r_i+1\,|\,\zeta=p\}}{\displaystyle\sum_{i=1}^{\infty}\frac{c_i(s_i+1)\binom{s_i}{r_i}}{(s_i+n+1)\binom{s_i+n}{r_i+k}}}. \tag{16}$$

If the number of components is small, formula (16) may be applied to practical computations, although these are more complicated than in the case of a simple Beta function.

6. Sampling acceptance based on the a priori Beta distribution

The most commonly used sampling acceptance plans do not take into account in their mathematical foundations the results of the past acceptance. The necessity of taking into consideration past results, however, is emphasized by special books with practical tendency (e. g. [4]). Usually this is done by some additional directives (e. g. occasional tightening or loosening.). In some cases a more accurate consideration of past results may be necessary. A well-known method for this purpose is that of the continuous sampling inspections (see e. g. [2]). In the existing continuous sampling procedures the choice of some of the parameters is quite arbitrary. On the contrary, in the method of a priori Beta distribution the determination of the parameters is based on the empirical data. The application of such a method was discussed by SKELLAM [17], CHAMPERNOWNE [4], BARNARD [1] and ODERFELD [9], [10], the most exhaustive treatment of the practical computations being given in [9]. Practical examples in which the approaching of the a priori distribution with that of Beta type turns out to be sufficient can be found in papers of ODERFELD [9] and HOPKINS [6].

In determining the parameters of the a priori Beta distribution ODERFELD uses the method of equating the moments. His formulae hold only under the supposition that the empirical data give the true fraction defectives of the preceding lots accepted, though in reality these are generally not available, because we know the fraction defectives of samples only. In addition, the examples given by him as well as the formulae adapted to practical computations show that he himself regarded the past empirical data to be sample results. In such cases the method can be applied only if generalized for the case of past data consisting of sample results.

This can be done, however, without difficulty. If the population fraction defective has a Beta distribution according to (14), then the probability of finding $\varkappa = k$ defectives in a sample of size n is ([17], [14])

$$\int_0^1 \frac{x^r(1-x)^{s-r}}{B(r+1, s-r+1)} \binom{n}{k} x^k (1-x)^{n-k}\, dx = \frac{(s+1)\binom{s}{r}\binom{n}{k}}{(s+n+1)\binom{s+n}{r+k}}.$$

As shown in [15], this distribution is of exceedance type.

The first two moments of x are

$$\mathbf{M}(x) = \frac{n(r+1)}{s+2},$$

$$\mathbf{M}(x^2) = \frac{n(n-1)(r+1)(r+2)}{(s+2)(s+3)} + \frac{n(r+1)}{s+2},$$

these can be equated with the empirical moments.

Performing the computations of ODERFELD's examples, according to the above modification, the difference will be considerable [14].

We wish to point out that the method can give absurd results (negative variance for the Beta distribution) if the number of empirical data used is not sufficiently large.

As the computation of the second moment is tedious, we can replace it by equating a quantile of prescribed level. In this case we have to solve the system of equations

$$p_0 = \frac{r+1}{s+2},$$

(17)

$$\beta = \int_0^{p_\beta} x^r (1-x)^{s-r} dx$$

for s and r. Here p_0 denotes the average fraction defective and p_β the quantile of level β (near to 1) of the empirical distribution of the fraction defective. The solution of this system of equations is very simple if the Poisson approximation is legitimate and may be performed by a special table or nomogram [14]. If p_β is based upon sample results, then the following correction is to be made for r:

$$r = r^* + \frac{(r^*+1)^2}{np_0}$$

where r denotes the corrected value and r^* the value given by (17). The correct value of s is to be computed by substituting the corrected value of r into the first equation of (17).

A more detailed treatment of the subject of this section can be found in the Hungarian version [14].

(Received 21 August 1959)

References

[1] G. A. Barnard, Sampling inspection and statistical decisions, *Journal Royal Stat. Soc.*, Ser. B, **16** (1954), pp. 151—165.

[2] A. H. Bowker, Continuous sampling plans, *Proc. 3rd Berkeley Symp. Math. Stat. Probability*, **5** (1956), pp. 75—85.

[3] G. Castelnuovo, *Calcolo delle probabilità* (Milano, 1919).

[4] D. G. Champernowne, The economics of sequential sampling procedure for defectives, *Appl. Statist.*, **2** (1953), pp. 118—130.

[5] H. A. Freeman, M. Friedman, F. Mosteller and W. A. Wallis, *Sampling inspection* (New York, 1948).

[6] J. W. Hopkins, An instance of negative hypergeometric sampling in practice. *Bull. Internat. Stat. Inst.*, **34** (1955), pp. 298-306.

[7] C. Jordan, Le théorème de probabilité de Poincaré, généralisé au cas de plusieurs variables indépendantes, *Acta Sci. Math. Szeged*, **7** (1934), pp. 103—111.

[8] E. C. Molina, Bayes' theorem: an expository presentation, *Bell System Technical Journal*, **10** (1931), pp. 273-287, and *Annals of Math. Stat.*, **2** (1931), pp. 23—37.

[9] J. Oderfeld, *Statystyczny odbior towarow klasyfikowanych wedlug alternatywy* (Warszawa, 1950).

[10] J. Oderfeld, On the dual aspect of sampling plans, *Coll. Math.*, **2** (1951), pp. 89—97.

[11] A. Rényi, On a new axiomatic theory of probability, *Acta Math. Acad. Sci. Hung.*, **6** (1955), pp. 285—335.

[12] В. И. Романовский, Теоремы двойственности для гипергеометрического распределения, Тр. Института Мат. Мех. Узб. ССР, **11** (1953), pp. 22—28.

[13] K. Sarkadi, On the rule of dualism concerning the Bayes' probability limits of the fraction defective, *MTA Alk. Mat. Int. Közl.*, **2** (1953), pp. 275—286.

[14] K. Sarkadi, On the a priori Beta distribution of fraction defective, *MTA Alk. Mat. Int. Közl.*, **2** (1953), pp. 287—293.

[15] K. Sarkadi, Generalized hypergeometric distributions, *Publ. Math. Inst. Hung. Acad. Sci.*, **2** (1957), pp. 59—69.

[16] L. E. Simon, *An engineers' manual of statistical methods* (New York, 1941).

[17] J. G. Skellam, A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials, *Journal Royal Stat. Soc.*, Ser. B, **10** (1948), pp. 257—261.

[18] H. Steinhaus, Quality control by sampling, *Coll. Math.*, **2** (1951), pp. 98—108.

[19] H. Steinhaus and S. Zubrzycki, On the comparison of two production processes and the rule of dualism, *Coll. Math.*, **5** (1958), pp. 103—115.

REMARK ON A THEOREM OF CINQUINI

By
A. SHARMA (Lucknow, India)
(*Presented by* P. TURÁN)

1. S. CINQUINI [1], [2], [3] and P. MONTEL [4], [5] have extended the mean value theorem for real variables into the field of complex variables. CINQUINI goes so far as to consider a system of number-pairs $(0, z_0)$, $(0, z_1)$, $(1, z_2)$ with the linear differential operator $\frac{d^2}{dz^2}$, but he puts this limitation on z_2 that it does not lie in a circle with $\frac{z_0+z_1}{2}$ as centre and $\left|\frac{z_1-z_0}{2}\right|$ as radius. In real variables, such a condition makes the set a conservative set in the terminology of BIRKHOFF. The object of this note is to use the method of CINQUINI to prove the following theorem wherein z_2 is taken to be $\frac{z_0+z_1}{2}$.

2. THEOREM. *If $f(z)$ is analytic function regular for $|z-\alpha|\leqq r$ and if $f''(\alpha)\neq 0$, then there exists a circle $|z-\alpha|=r_0'<r$ such that if z_0 and z_1 are any two points inside the circle satisfying the conditions that $|z_0-\alpha|=$ $=|z_1-\alpha|=r_0<r_0'$ and if $\alpha=\frac{z_0+z_1}{2}$, then*

$$\begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ f(z_0) & f(z_1) & f'\left(\frac{z_0+z_1}{2}\right) \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ z_0^3 & z_1^3 & 3\left(\frac{z_0+z_1}{2}\right)^2 \end{vmatrix} \frac{f'''(\xi)}{3!}$$

where ξ is a point inside the circle $|z-\alpha|=r_0'$.

Let the power series for $f(z)$ in the neighbourhood of $z=\alpha$ be

$$f(z)=\sum_{\nu=0}^{\infty} a_\nu (z-\alpha)^\nu. \tag{2. 1}$$

Let M be the maximum modulus of $f(z)$ on $|z-\alpha|=R$, where R is less than the radius of convergence of the power series (2. 1), then

$$|a_\nu|\leqq \frac{M}{R^\nu} \qquad (\nu=0, 1, 2, \ldots).$$

We have on using (2.1)

$$\begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ f(z_0) & f(z_1) & f'(\alpha) \end{vmatrix} = -\sum_{\nu=2}^{\infty} a_\nu[(z_1-\alpha)^\nu-(z_0-\alpha)^\nu].$$

Also

$$\begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ z_0^3 & z_1^3 & 3\alpha^2 \end{vmatrix} = -\frac{(z_1-z_0)^3}{4}.$$

Then

$$Q_2(z_0, z_1) = \begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ f(z_0) & f(z_1) & f'(\alpha) \end{vmatrix} \div \begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ z_0^3 & z_1^3 & 3\alpha^2 \end{vmatrix} =$$

$$= \frac{4\sum_{\nu=2}^{\infty} a_\nu[(z_1-\alpha)^\nu-(z_0-\alpha)^\nu]}{[(z_1-\alpha)-(z_0-\alpha)]^3} =$$

$$= \frac{4\sum_{\nu=2}^{\infty} a_\nu[(z_1-\alpha)^{\nu-1}+(z_1-\alpha)^{\nu-2}(z_0-\alpha)+\cdots+(z_0-\alpha)^{\nu-1}]}{[(z_1-\alpha)-(z_0-\alpha)]^2}.$$

Since

$$z_1-\alpha+z_0-\alpha=0,$$

$$(z_0-\alpha)^2+(z_1-\alpha)(z_0-\alpha)+(z_1-\alpha)^2=\left(\frac{z_1-z_2}{2}\right)^2$$

and

$$(z_0-\alpha)^3+(z_0-\alpha)^2(z_1-\alpha)+(z_0-\alpha)(z_1-\alpha)^2+(z_1-\alpha)^3=0,$$

we have

$$Q_2(z_0, z_1) = a_3+\sum_{\nu=5}^{\infty} a_\nu F_\nu$$

where

$$F_\nu = 4\,\frac{(z_1-\alpha)^{\nu-1}+(z_1-\alpha)^{\nu-2}(z_0-\alpha)+\cdots+(z_0-\alpha)^{\nu-1}}{[(z_1-\alpha)-(z_0-\alpha)]^2}.$$

Also

$$\frac{f'''(z)}{3!} = a_3+\sum_{\nu=4}^{\infty}\binom{\nu}{3} a_\nu(z-\alpha)^{\nu-3}.$$

Let $\mu(\alpha; r_0)$ be the maximum modulus of the function

$$\psi(z) = \sum_{\nu=4}^{\infty}\binom{\nu}{3} a_\nu(z-\alpha)^{\nu-3}$$

on the circumference of $|z-\alpha|=r_0$. Since $\psi(\alpha)=0$, $\psi(z)$ will assume inside the circumference at least once every value less than $\mu(\alpha; r_0)$. It is then enough to take r_0 in such a way that the maximum modulus of $\sum_{\nu=5}^{\infty} a_\nu F_\nu$ shall be less than $\mu(\alpha; r_0)$.

For ν even, $F_\nu=0$. For ν odd,

$$F_\nu=\frac{(z-\alpha)^{\nu-1}}{[(z_1-\alpha)-(z_0-\alpha)]^2}=\frac{[(z_1-\alpha)-(z_0-\alpha)]^{\nu-3}}{2^{\nu-1}}.$$

Then

$$\left|\sum_{\nu=5}^{\infty} a_\nu F_\nu\right| \leqq \sum_{m=2}^{\infty}\frac{M}{R^{2m+1}}\frac{(2r_0)^{2m-2}}{2^{2m}} \leqq \frac{M}{4R^3}\frac{r_0^2}{R^2-r_0^2}<\frac{M}{4R^4}\frac{r_0^2}{R-r_0}.$$

Now

$$\mu(\alpha;\, r_0)\geqq 4|a_4|\,r_0-\left|\sum_{\nu=5}^{\infty}\binom{\nu}{3}\frac{M}{R^\nu}r_0^{\nu-3}\right|\geqq 4|a_4|\,r_0-\frac{Mr_0^2}{R^5}\left|\sum_{\nu=5}^{\infty}\binom{\nu}{3}\left(\frac{r_0}{R}\right)^{\nu-5}\right|.$$

Since

$$\sum_{\nu=5}^{\infty}\binom{\nu}{3}\left(\frac{r_0}{R}\right)^{\nu-5}=\frac{R(10R^3-20R^2r_0+15Rr_0^2-4r_0^3)}{(R-r_0)^4}<\frac{10R^4}{(R-r_0)^4},$$

we have

$$\mu(\alpha; r_0)>4|a_4|\,r_0-\frac{10Mr_0^2}{(R-r_0)^5}.$$

Then the relation

$$\sum_{\nu=4}^{\infty}\binom{\nu}{3}a_\nu(\xi-\alpha)^{\nu-3}=\sum_{\nu=5}^{\infty}a_\nu F_\nu$$

is surely satisfied for some ξ if

$$\frac{Mr_0^2}{4R^4(R-r_0)}<4|a_4|\,r_0-\binom{5}{3}\frac{Mr_0^2}{(R-r_0)^5},$$

i. e. if

$$\frac{Mr_0(41R^4-4R^3r_0+6R^2r_0^2-4Rr_0^3+r_0^4)}{4R^4(R-r_0)^5}<4|a_4|.$$

As the first member of the last inequality tends to zero for $r_0\to 0$, while the second does not depend upon r_0 and is always positive, a value r_0' of r_0 certainly exists for which this inequality is satisfied, and as every $r_0<r_0'$ satisfies this inequality, the theorem stated is proved.

We also obtain as CINQUINI has done the effective value of r_0'. Put

$$\frac{16 R^4 |a_4|}{M} = \Lambda$$

and consider

$$\Phi(r_0) = \Lambda (R - r_0)^5 - (41 R^4 r_0 - 4 R^3 r_0^2 + 6 R^2 r_0^3 - 4 R r_0^4 + r_0^5)$$

and its first two derivatives

$$\Phi'(r_0) = -5 \Lambda (R - r_0)^4 - (41 R^4 - 8 R^3 r_0 + 18 R^2 r_0^2 - 16 R r_0^3 + 5 r_0^4),$$

$$\Phi''(r_0) = 20 \Lambda (R - r_0)^3 + (8 R^3 - 36 R^2 r_0 + 48 R r_0^2 - 20 r_0^3).$$

As $\Phi''(r_0) > 0$ for $0 < r_0 \leqq \frac{2R}{9}$, it follows that for this value of r_0

$$\Phi(r_0) > \Phi(0) + r_0 \Phi'(0).$$

Hence

$$\Phi(0) + r_0 \Phi'(0) \equiv \Lambda R^5 + r_0(-5 \Lambda R^4 - 41 R^4) \geqq 0$$

if

$$r_0 \leqq \frac{\Lambda R}{5\Lambda + 41}.$$

We may then take

$$r_0' = \frac{R}{5 + \dfrac{41}{16 |a_4| R^4}}.$$

(Received 1 September 1959)

References

[1] S. CINQUINI, Sopra una formula di curtiss, *Annali di Math.*, Ser. 4, **6** (1933—34), pp. 153—161.
[2] S. CINQUINI, Sopra un estensione di una formula di curtiss, *Rend. Ist. Lomb.*, **70** (1937).
[3] S. CINQUINI, Sopra una nuova estensione di una formula di curtiss, *Rend. Circ. Mat. Palermo*, **61** (1937), pp. 73—83.
[4] P. MONTEL, Sur quelques propriétés des différences devisées, *Journ. de Math.*, Ser. 4, **16** (1937), pp. 219—231.
[5] P. MONTEL, Sur une formula de Darboux et les polynômes d'interpolation, *Annali. Pisa.*, Ser. 2, **1** (1932), pp. 371—384.

ON THE CENTRAL LIMIT THEOREM FOR THE SUM OF A RANDOM NUMBER OF INDEPENDENT RANDOM VARIABLES

By

A. RÉNYI (Budapest), member of the Academy

§ 1. Introduction

Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ denote throughout the present paper a sequence of independent and identically distributed random variables with mean value 0 and variance 1. Let us put

$$\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n \tag{1}$$

and

$$\eta_n = \frac{\zeta_n}{\sqrt{n}}. \tag{2}$$

Then by the simplest case of the central limit theorem

$$\lim_{n \to +\infty} \mathbf{P}(\eta_n < x) = \Phi(x) \qquad (-\infty < x < +\infty). \tag{3}$$

Here and in what follows $\mathbf{P}(\ldots)$ denotes the probability of the event in the brackets and $\Phi(x)$ the standard form of the normal distribution function, i. e.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du. \tag{4}$$

In what follows we shall investigate the limiting distribution of the random variables η_{ν_n} for $n \to +\infty$ where ν_n $(n = 1, 2, \ldots)$ is a sequence of positive integer-valued random variables. The importance of this question, i. e. of the investigation of the behaviour of the sum of a random number of random variables and the role played by such sums in sequential analysis, in random walk problems, in connection with Monte Carlo methods and in other chapters of probability theory is nowadays well known.

It is easy to see that if the positive integer-valued random variable ν_n is for each $n = 1, 2, \ldots$ independent of all the random variables ξ_k $(k = 1, 2, \ldots)$ and if ν_n tends in probability to $+\infty$ (i. e if $\lim_{n \to +\infty} \mathbf{P}(\nu_n > c) = 1$ for any $c > 0$), then (3) remains valid if we replace n by ν_n in it, i.e. we have

$$\lim_{n \to +\infty} \mathbf{P}(\eta_{\nu_n} < x) = \Phi(x). \tag{5}$$

This can be shown as follows:

By virtue of the supposition[1] of the independence of ν_n and the variables ξ_k, we have

$$\mathbf{P}(\eta_{\nu_n} < x,\ \nu_n = k) = \mathbf{P}(\eta_k < x)\mathbf{P}(\nu_n = k). \tag{6}$$

It follows from (6) that putting

$$p_{nk} = \mathbf{P}(\nu_n = k) \tag{7}$$

we have

$$\mathbf{P}(\eta_{\nu_n} < x) = \sum_{k=1}^{\infty} \mathbf{P}(\eta_k < x) p_{nk}. \tag{8}$$

But by our suppositions (p_{nk}) is the matrix of a regular method of summation[2] and thus (3) and (8) imply (5).

If it is not supposed that ν_n is independent of the ξ_k more stringent conditions on the convergence of ν_n towards infinity are needed.

It follows from a more general result of F. J. ANSCOMBE [1] that the following theorem holds:

THEOREM 1. *If ν_n is such a sequence of positive integer-valued random variables that $\frac{\nu_n}{n}$ converges for $n \to +\infty$ in probability to a constant $c > 0$, then* (5) *holds.*

A straightforward proof of Theorem 1 has been given in [2].[3] In the present paper we generalize Theorem 1 in that we investigate the case when $\frac{\nu_n}{n}$ tends in probability to a positive random variable λ, instead of a constant. Similarly as in Theorem 1, nothing is supposed about the dependence of ν_n on the random variables ξ_k.

Thus we prove the following

THEOREM 2. *If ν_n $(n = 1, 2, \ldots)$ is such a sequence of positive integer-valued random variables that $\frac{\nu_n}{n}$ converges in probability to a positive random variable λ having a discrete distribution, then* (5) *holds.*

[1] We need here not the full strength of the supposition that the random variables ν_n and ξ_k are independent, only that the events $\nu_n = k$ and $\eta_k < x$ should be independent.

[2] See e. g. [7], Ch. III, Theorem 2.

[3] K. L. CHUNG kindly called my attention to the fact that the main idea of the proof (the application of the inequality of KOLMOGOROV (see Lemma 3 below)) given in [2] is due to DOEBLIN (see [3]). It has been recently proved by J. MOGYORÓDI (oral communication) that the method in question can lead to the proof of the most general form of ANSCOMBE's theorem.

§ 2. Proof of Theorem 2

We shall need five lemmas. Lemma 1 is due to H. CRAMÉR (see [4], p. 254); Lemma 2 is almost trivial; its proof is given below; Lemma 3 is the well known inequality of KOLMOGOROV; Lemma 4 is proved in the paper [5] of the author. Lemma 5 is a special case of Theorem 2 which is an easy consequence of Lemma 4 and which is utilized in the proof of the general Theorem. In what follows $\vartheta_n \Rightarrow \vartheta$ means that ϑ_n converges in probability to ϑ, i. e. $\lim_{n\to+\infty} \mathbf{P}(|\vartheta_n - \vartheta| > \varepsilon) = 0$ for all $\varepsilon > 0$.

LEMMA 1. *If ϑ_n is a sequence of random variables having a limiting distribution,*[4] *ϱ_n and δ_n are random variables such that $\varrho_n \Rightarrow 1$ and $\delta_n \Rightarrow 0$, then the sequence $\varrho_n \vartheta_n + \delta_n$ has the same limiting distribution as the sequence ϑ_n.*

LEMMA 2. *If the distribution of the random variables ϑ_n tends to a limiting distribution and $\delta_n \Rightarrow 0$, then $\vartheta_n \delta_n \Rightarrow 0$.*

PROOF OF LEMMA 2. We have evidently for any $\varepsilon > 0$ and $N > 0$

$$\mathbf{P}(|\vartheta_n \delta_n| > \varepsilon) \leqq \mathbf{P}(|\vartheta_n| > N) + \mathbf{P}\left(|\delta_n| > \frac{\varepsilon}{N}\right). \tag{9}$$

By choosing N sufficiently large, the limit of the right-hand side of (9) can be made arbitrarily small, which proves our assertion.

LEMMA 3. (The inequality of KOLMOGOROV.) *Let $\tau_1, \tau_2, \ldots, \tau_n$ be independent random variables with mean 0 and variance 1, then for any $c > 1$*

$$\mathbf{P}\left(\max_{1 \leqq k \leqq n} |\tau_1 + \tau_2 + \cdots + \tau_k| \geqq c\sqrt{n}\right) \leqq \frac{1}{c^2}.$$

LEMMA 4. *If τ_n $(n = 1, 2, \ldots)$ is a sequence of independent random variables such that putting*

$$\sigma_n = \frac{1}{B_n} \sum_{k=1}^{n} \tau_k \quad \textit{where} \quad B_n \to +\infty$$

the distribution of the random variable σ_n tends to a limiting distribution, then the conditional distribution of σ_n under any condition having a positive probability, tends to the same limiting distribution.

[4] By saying that a sequence ϑ_n of random variables has a limiting distribution we mean throughout the paper that there exists a distribution function $F(x)$ such that $\mathbf{P}(\vartheta_n < x) \to F(x)$ for every x which is a point of continuity of $F(x)$.

We shall prove Theorem 2 by proving first that it is valid in the special case when $\nu_n = [n\lambda]$ where λ is a positive random variable having a discrete distribution (here and in what follows $[x]$ denotes the integral part of the real number x), and then reduce the general case to this special case. Thus we prove first the following

LEMMA 5. *If* $\nu_n = [n\lambda]$ *where* λ *is a positive random variable having a discrete distribution, then*

$$\lim_{n\to+\infty} \mathbf{P}(\eta_{\nu_n} < x) = \Phi(x).$$

PROOF OF LEMMA 5. Let (here and in what follows) l_k $(0 < l_1 < l_2 < \cdots)$ denote the values taken on by λ with positive probability (the sequence l_k may be finite or infinite); let A_k denote the event that $\lambda = l_k$. Then we have, evidently, denoting by $\mathbf{P}(A|B)$ the conditional probability of the event A under condition B,

$$\mathbf{P}(\eta_{\nu_n} < x) = \sum_{k=1}^{\infty} \mathbf{P}(\eta_{[nl_k]} < x \,|\, A_k)\,\mathbf{P}(A_k).$$

As by Lemma 4

$$\lim_{n\to+\infty} \mathbf{P}(\eta_{[nl_k]} < x \,|\, A_k) = \Phi(x) \quad \text{for} \quad k = 1, 2, \ldots,$$

the assertion of Lemma 5 follows immediately.

Now we turn to the proof of Theorem 2.

We have clearly

$$\eta_{\nu_n} = \eta_{[\lambda n]} + \sqrt{\frac{[\lambda n]}{\nu_n}}\left(\frac{\zeta_{\nu_n} - \zeta_{[\lambda n]}}{\sqrt{[\lambda n]}}\right) + \frac{\zeta_{[\lambda n]}}{\sqrt{[\lambda n]}}\left(\sqrt{\frac{[\lambda n]}{\nu_n}} - 1\right). \tag{10}$$

As evidently, $\dfrac{\nu_n}{[\lambda n]} \Rightarrow 1$, the third term tends in probability to 0 by Lemma 2. Thus by Lemma 1 and Lemma 5 to prove Theorem 2 it suffices to show that

$$\frac{\zeta_{\nu_n} - \zeta_{[\lambda n]}}{\sqrt{[\lambda n]}} \Rightarrow 0. \tag{11}$$

This can be done as follows: Let us denote by $B_n(\varrho)$, where $\varrho > 0$, the event $|\nu_n - [\lambda n]| < n\varrho$ and by $\overline{B_n(\varrho)}$ the contrary event, and put $n_k = [nl_k]$; let us choose an arbitrary $\varepsilon > 0$ and let C_{nk} denote the event $\dfrac{|\zeta_{\nu_n} - \zeta_{n_k}|}{\sqrt{n_k}} > \varepsilon$; then we have

$$\mathbf{P}\left(\left|\frac{\zeta_{\nu_n} - \zeta_{[\lambda n]}}{\sqrt{[\lambda n]}}\right| > \varepsilon\right) \leqq \sum_{k=1}^{\infty} \mathbf{P}(A_k B_n(\varrho) C_{nk}) + \mathbf{P}(\overline{B_n(\varrho)}). \tag{12}$$

(Here and in what follows the product of two or more events denotes the event consisting in the joint occurrence of these events.) It follows by Lemma 3 that

$$\text{(13)} \qquad \mathbf{P}(A_k B_n(\varrho) C_{nk}) \leqq \mathbf{P}\left(\max_{|l-n_k| \leqq \varrho n} \frac{|\zeta_l - \zeta_{n_k}|}{\sqrt{n_k}} > \varepsilon\right) \leqq \frac{2\varrho}{\varepsilon^2 l_k}.$$

Thus if D_M denotes event $\lambda \geqq l_M$, it follows that

$$\text{(14)} \qquad \mathbf{P}\left(\left|\frac{\zeta_{\nu_n} - \zeta_{[\lambda n]}}{\sqrt{[\lambda n]}}\right| > \varepsilon\right) \leqq \mathbf{P}(D_M) + \frac{2\varrho}{\varepsilon^2} \sum_{k=1}^{M-1} \frac{1}{l_k} + \mathbf{P}(\overline{B_n(\varrho)}).$$

Let now $\delta > 0$ be arbitrarily small; let us choose first M so large that $\mathbf{P}(D_M) < \frac{\delta}{3}$; if M has been fixed in this way, we choose ϱ so small that

$$\frac{2\varrho}{\varepsilon^2} \sum_{k=1}^{M-1} \frac{1}{l_k} < \frac{\delta}{3}.$$

After this we choose $n_0 = n_0(\varepsilon, \delta)$ so large that for $n \geqq n_0$ we should have $\mathbf{P}(\overline{B_n(\varrho)}) < \frac{\delta}{3}$. It follows from (14) that

$$\text{(15)} \qquad \mathbf{P}\left(\frac{|\zeta_{\nu_n} - \zeta_{[\lambda n]}|}{\sqrt{[\lambda n]}} > \varepsilon\right) \leqq \delta \quad \text{if} \quad n \geqq n_0.$$

As ε and δ may both be chosen arbitrarily small, it follows that (11) holds; thus Theorem 2 is proved.

§ 3. Some additional remarks

We add some remarks on possible generalizations of Theorem 2. Clearly, the above method can be applied to the case also when the random variables ξ_k are not identically distributed, but are such that the distribution of η_n tends to a limiting distribution. We do not go, however, into details.

It seems likely that the assertion of Theorem 2 holds also without the restriction that the positive random variable λ has a distribution of the *discrete* type.

It is clear that supposition $\frac{\nu_n}{n} \Rightarrow \lambda$ in Theorem 2 can be replaced by the more general supposition $\frac{\nu_n}{\omega(n)} \Rightarrow \lambda$ where $\omega(n)$ is an arbitrary positive function tending to $+\infty$ for $n \to +\infty$.

It can be further shown easily that the weaker supposition that the distribution of $\frac{\nu_n}{n}$ tends to that of λ is *not* sufficient to ensure the validity of (5). As a matter of fact, put $\nu_n = n$ if $\eta_n \geqq 0$ and $\nu_n = 2n$ if $\eta_n < 0$. In this case

$$\lim_{n\to+\infty} \mathbf{P}\left(\frac{\nu_n}{n} = 1\right) = \lim_{n\to+\infty} \mathbf{P}\left(\frac{\nu_n}{n} = 2\right) = \frac{1}{2},$$

but

$$\lim_{n\to+\infty} \mathbf{P}(\eta_n < 0,\ \eta_{2n} \geqq 0) = \int_0^\infty (1 - \Phi(x))\, d\Phi(x) = \frac{1}{8}$$

and thus we obtain

$$\lim_{n\to+\infty} \mathbf{P}(\eta_{\nu_n} < 0) = \frac{3}{8}.$$

But this means that (5) does not hold, as $\Phi(0) = \frac{1}{2} \neq \frac{3}{8}$.

(Received 3 September 1959)

References

[1] F. J. Anscombe, Large-sample theory of sequential estimation, *Proc. Cambridge Phil. Soc.*, **48** (1952), p. 600.

[2] A. Rényi, On the asymptotic distribution of the sum of a random number of independent random variables, *Acta. Math. Acad. Sci. Hung.*, **8** (1957), pp. 193—199.

[3] W. Doeblin, Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables, *Bull. Soc. Math. France*, **66** (1938), pp. 210—220.

[4] H. Cramér, *Mathematical methods of statistics* (Princeton, 1956).

[5] A. Rényi, On mixing sequences of sets, *Acta. Math. Acad. Sci. Hung.*, **9** (1958), pp. 215—228.

[6] A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Math. u. Grenzgebiete (Berlin, 1933).

[7] G. H. Hardy, *Divergent series* (Oxford, 1949).

UNGLEICHUNGEN FÜR UMFANG, FLÄCHENINHALT UND TRÄGHEITSMOMENT KONVEXER KURVEN

Von

HORST SACHS (Halle)

(*Vorgelegt von* G. HAJÓS)

Inhalt

§ 1. Einleitung; Resultate

Den Anlaß zu dieser Arbeit gab folgender mir aus der Literatur nicht bekannter

SATZ. *Für jede ebene konvexe Figur mit der Randkurve $\mathfrak{C}$ der Länge 2π, welche den Ursprung als inneren Punkt enthält, gilt*

$$\int_{\mathfrak{C}} r^2 ds \geqq \int_{\mathfrak{C}} r^2 d\varphi \tag{1}$$

(r, φ: Polarkoordinaten), und das Gleichheitszeichen steht genau dann, wenn $\mathfrak{C}$ der Einheitskreis mit dem Ursprung als Mittelpunkt ist.

Beim Beweise ergab sich eine Reihe von Ungleichungen für die in der Überschrift genannten elementaren Größen, die im folgenden systematisch hergeleitet werden sollen; sie scheinen mir auch deshalb von Interesse, weil dadurch auf neue Weise die klassische isoperimetrische Ungleichung (in verschärfter Form) in ein tiefer liegendes System von Ungleichungen eingeordnet wird.

Zuerst werden Raumkurven, dann ebene und schließlich konvexe Kurven behandelt; dabei gewinnt man eine Übersicht über den Gültigkeitsbereich der betrachteten Ungleichungen. Das Hauptresultat dieser Arbeit ist die Ungleichung

$$8\pi^2 LI - L^4 - (4\pi F)^2 \geqq 0; \tag{14}$$

Hauptbeweismittel ist die Schwarzsche Ungleichung.

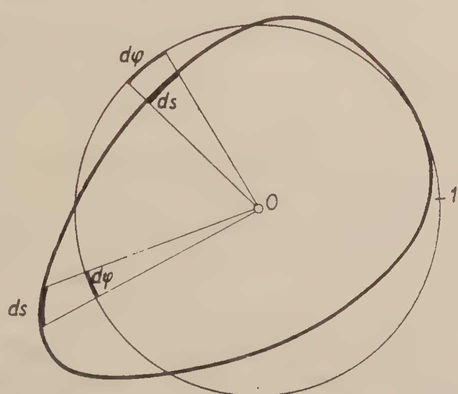


Fig 1

$\mathfrak{C}$ sei eine rektifizierbare geschlossene Kurve der Länge L, I bezeichne das minimale (das ist das auf den Schwerpunkt S von $\mathfrak{C}$ bezogene) Trägheitsmoment von $\mathfrak{C}$:

$$I = \int_{\substack{\mathfrak{C} \\ O=S}} \mathfrak{r}^2 ds,$$

und F sei, falls $\mathfrak{C}$ eine ebene Kurve ist, der von $\mathfrak{C}$ umschlossene Flächeninhalt. Wir führen folgende Bezeichnungen ein:

$\mathfrak{K}$ Kreis,

$\mathfrak{D}$ gleichseitiges Dreieck,

$\mathfrak{N}$ Nadel (das ist die aus den beiden Ufern und den Endpunkten einer geraden Strecke bestehende Kurve).

Dann besteht für *beliebige* $\mathfrak{C}$ des n-dimensionalen euklidischen Raumes die Ungleichung

$$L^3 - 4\pi^2 I \underset{\mathfrak{K}}{\geqq} 0, \tag{2}$$

wo das Zeichen $\underset{\mathfrak{K}}{\geqq}$ bedeuten soll, daß der Fall der Gleichheit genau dann eintritt, wenn $\mathfrak{C}$ ein Kreis ist.[1]

Für *ebene* $\mathfrak{C}$ beweisen wir eine weitere Ungleichung: Wir führen als neuen Parameter an Stelle der Bogenlänge s für die Punkte von $\mathfrak{C}$ die Größe $\Phi = \frac{2\pi}{L} s$ ein, so daß

$$\left(\frac{dx}{d\Phi}\right)^2 + \left(\frac{dy}{d\Phi}\right)^2 = \frac{L^2}{4\pi^2}.$$

Dann ist

$$L^3 = 2\pi L \int_0^{2\pi} \left\{\left(\frac{dy}{d\Phi}\right)^2 + \left(\frac{dx}{d\Phi}\right)^2\right\} d\Phi,$$

$$-8\pi L F = 2\pi L \int_0^{2\pi} \left\{-2x\frac{dy}{d\Phi} + 2y\frac{dx}{d\Phi}\right\} d\Phi,$$

$$4\pi^2 I = 2\pi L \int_0^{2\pi} \{x^2 + y^2\}\, d\Phi.$$

[1] Siehe [6], Satz 3; (2) ist im wesentlichen äquivalent einem Lemma von WIRTINGER (siehe [1], S. 105, sowie [2], S. 184 ff) und wird am schnellsten bewiesen, wenn man $L = 2\pi$ setzt und $\mathfrak{r}(s)$ in eine Fourierreihe entwickelt (man beachte die Verwandtschaft mit dem bekannten Hurwitzschen Beweis [3] der klassischen isoperimetrischen Ungleichung).

und durch Addition ergibt sich

$$L^3-8\pi LF+4\pi^2 I=2\pi L\int\limits_0^{2\pi}\left\{\left(x-\frac{dy}{d\Phi}\right)^2+\left(y+\frac{dx}{d\Phi}\right)^2\right\}d\Phi, \tag{3}$$

folglich

$$L^3-8\pi LF+4\pi^2 I\underset{\mathfrak{K}}{\geqq}0, \tag{4}$$

denn die rechte Seite von (3) verschwindet nur dann, wenn längs $\mathfrak{C}$ $x=\frac{dy}{d\Phi}$ und $y=-\frac{dx}{d\Phi}$ ist, d. h. wenn $\mathfrak{C}$ ein Kreis mit dem Ursprung als Mittelpunkt ist.

Durch Addition ergibt sich aus (2) und (4) sofort die klassische isoperimetrische Ungleichung

$$L^2-4\pi F\underset{\mathfrak{K}}{\geqq}0. \tag{5}$$

Wir bekommen hier auch eine erste Verschärfung derselben, indem wir (2) in der Form

$$2L(L^2-4\pi F)\underset{\mathfrak{K}}{\geqq}L^3-8\pi LF+4\pi^2 I \tag{6}$$

schreiben und (4) beachten; auf eine geometrische Deutung des rechten Ausdrucks $L^3-8\pi LF+4\pi^2 I$ kommen wir noch zurück.

Aus (2) folgern wir noch

$$L(L^2-4\pi F)\underset{\mathfrak{K}}{\geqq}4\pi(\pi I-LF) \tag{7}$$

und aus (4)

$$4\pi(\pi I-LF)\underset{\mathfrak{K}}{\geqq}-L(L^2-4\pi F); \tag{8}$$

(7) und (8) zusammen ergeben eine andere Verschärfung von (5):

$$L^2-4\pi F\underset{\mathfrak{K}}{\geqq}\frac{4\pi}{L}|\pi I-LF|. \tag{9}$$

Wir bemerken, daß es Kurven gibt mit $\pi I-LF>0$ (siehe (16)) als auch solche (nichtkonvexe) mit $\pi I-LF<0$.[2]

Bezeichnen wir mit Q das Quadratmittel der paarweisen Entfernung der Punkte von $\mathfrak{C}$:

$$Q=\left\{\frac{1}{L^2}\int\limits_{\mathfrak{C}}\int\limits_{\mathfrak{C}}|PP'|^2 ds\,ds'\right\}^{\frac{1}{2}},$$

[2] Zur Konstruktion von Kurven mit $\pi I-LF<0$ vergleiche [9], S. 363.

so besteht zwischen Q, I und L die Relation

$$LQ^2 = 2I,\tag{10}$$ [3]

und wir gewinnen aus (2) und (9) die Ungleichungen

$$\left\{\begin{array}{l} L^2 - 2\pi^2 Q^2 \underset{\mathfrak{K}}{\geqq} 0, \\ L - \sqrt{2}\pi Q \underset{\mathfrak{K}}{\geqq} 0, \end{array}\right.\tag{2'}$$

$$L^2 - 4\pi F \underset{\mathfrak{K}}{\geqq} 2\pi |\pi Q^2 - 2F|.\tag{9'}$$

Die Ungleichungen (2), (2') und (5) lassen sich zu folgendem (bekannten) Satz zusammenfassen:

Unter allen geschlossenen rektifizierbaren Kurven gleicher Länge liefert der Kreis und nur dieser den größten Wert für das minimale Trägheitsmoment,[4] *das Quadratmittel der Entfernung*[4] *und den umschlossenen Flächeninhalt.*[5]

Weiter gilt für beliebige ebene Kurven $\mathfrak{C}$ die Ungleichung

$$LI - 4F^2 \underset{\mathfrak{K}}{\geqq} 0.\tag{10a}$$ [5a]

Wir kommen zurück zur Ungleichung (4). Das Integral auf der rechten Seite von (3) — bei beliebiger Lage des Ursprungs — ist gleich

$$\int_0^{2\pi} \mathbf{d}^2 d\Phi = 2\pi \frac{1}{L} \int_{\mathfrak{C}} \mathbf{d}^2 ds\tag{11}$$

mit $\mathbf{d} = \mathbf{r} - \frac{L}{2\pi}\mathbf{n}$, wo $\mathbf{n} = \left\{\frac{dy}{ds}, -\frac{dx}{ds}\right\}$ den Einheitsvektor in Richtung

[3] Siehe [6], Bemerkung (1) (S. 124; es ist $Q = M_2[\mathfrak{C}]$); die entsprechende Gleichung für endliche Summen findet sich schon bei Steiner. — (10) gilt auch für Raumkurven.

[4] Siehe [6], Sätze 1 und 3.

[5] Hier beschränken wir uns auf ebene Vergleichskurven.

[5a] Zum Beweis siehe [5], S. 49, Ungleichung (17); dort findet sich auch der Hinweis, daß (10a) leicht aus der Schwarzschen Ungleichung folgt:

$$4F^2 = \left\{\int_{\mathfrak{C}} (xy' - x'y)ds\right\}^2 \leqq \int_{\mathfrak{C}} (x^2 + y^2)ds \cdot \int_{\mathfrak{C}} (x'^2 + y'^2)ds = IL.$$

Man beachte auch [5], S. 50, wo für konvexe $\mathfrak{C}$ eine Ungleichung

$$\varkappa IL \leqq 4F^2$$

bewiesen wird ($\varkappa$ hat eine einfache geometrische Bedeutung).

der äußeren Normalen bezeichnet. Für den Kreis mit dem Ursprung als Mittelpunkt und nur für diesen ist $\mathbf{d} \equiv \mathbf{o}$, das Integral (11) gibt uns also ein Maß für die Abweichung der Kurve $\mathfrak{C}$ von der Kreisgestalt, bezogen auf den Ursprung. Man zeigt leicht, daß das Integral $\frac{1}{L}\int\limits_{\mathfrak{C}} \mathbf{d}^2 ds$ bei variabler Lage des Ursprungs seinen kleinsten Wert genau dann annimmt, wenn der Ursprung mit dem Schwerpunkt von $\mathfrak{C}$ zusammenfällt. Die (nichtnegative) Quadratwurzel aus diesem Minimum bezeichnen wir mit R:

$$R = \left\{ \frac{1}{L} \int\limits_{\substack{\mathfrak{C} \\ O=S}} \mathbf{d}^2\, ds \right\}^{\frac{1}{2}};$$

nach (3) und (10) gilt

$$4\pi^2 R^2 = \frac{1}{L}(L^3 - 8\pi L F + $$
$$+ 4\pi^2 I) = L^2 + 2\pi^2 Q^2 - 8\pi F.$$

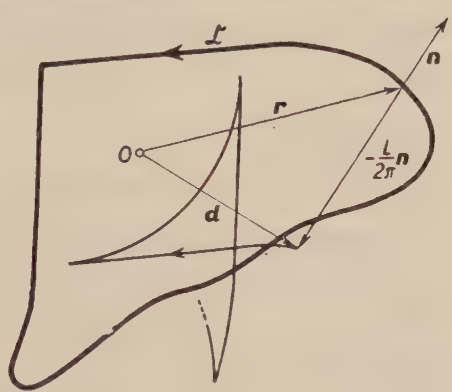


Fig. 2

Damit ist dem Ausdruck $L^3 - 8\pi L F + 4\pi^2 I$ eine einfache geometrische Bedeutung beigelegt. Jetzt können wir auch die Ungleichung (6) folgendermaßen schreiben:

$$(6') \qquad L^2 - 4\pi F \underset{\mathfrak{K}}{\geqq} 2\pi^2 R^2.$$ [6]

Auch die linke Seite von (6′) hat für (rektifizierbare) Jordankurven eine einfache geometrische Bedeutung, welche sich am leichtesten angeben läßt, wenn wir wieder den Vektor $\mathbf{d}$ zu Hilfe nehmen: Das "isoperimetrische Defizit" $\Delta = \frac{1}{4\pi}(L^2 - 4\pi F)$ ist bis auf das Vorzeichen gleich der vom Vektor $\mathbf{d}$

[6] (6′) läßt sich auf anderem Wege weiter verschärfen zu

$$(6'') \qquad L^2 - 4\pi F \geqq 2\pi^2 (R^2 + \sqrt{S^4 - T^4})$$

mit

$$S^4 = \frac{1}{L^2} \int\limits_{\mathfrak{C}} \int\limits_{\mathfrak{C}} (\mathbf{d}, \mathbf{d}')^2\, ds\, ds', \qquad T^4 = \frac{1}{L^2} \int\limits_{\mathfrak{C}} \int\limits_{\mathfrak{C}} [\mathbf{d}, \mathbf{d}']^2\, ds\, ds'$$

(die Punkte P, P' durchlaufen unabhängig voneinander die Kurve $\mathfrak{C}$; Ursprung ist der Schwerpunkt); es gilt $S^4 - T^4 \geqq 0$, $S^4 + T^4 = R^4$. — Für Jordankurven darf in (6″) das Zeichen $\geqq$ durch $\underset{\mathfrak{K}}{\geqq}$ ersetzt werden, aber für die Nadel z. B. gilt das Gleichheitszeichen.

überstrichenen Fläche, wenn der Punkt P (Fig. 2) die Kurve $\mathfrak{C}$ einmal im positiven Sinne durchläuft.[7]

*

Gehen wir zu *konvexen* Vergleichskurven über, so finden wir drei weitere Ungleichungen für L, F und I: zunächst die triviale

$$F \underset{\mathfrak{K}}{\geqq} 0,\text{[8]} \tag{12}$$

ferner, als Gegenstück zu (2),

$$54I - L^3 \underset{\mathfrak{T}}{\geqq} 0\text{[9]} \tag{13}$$

und schließlich

$$8\pi^2 LI - L^4 - (4\pi F)^2 \underset{\mathfrak{K}}{\geqq} 0.\text{[10]} \tag{14}$$

Wegen (10) schreiben wir statt (13) auch

$$27Q^2 - L^2 \underset{\mathfrak{T}}{\geqq} 0, \quad 3\sqrt{3}Q - L \underset{\mathfrak{T}}{\geqq} 0, \tag{13'}$$

und statt (14)

$$(2\pi LQ)^2 - L^4 - (4\pi F)^2 \underset{\mathfrak{K}}{\geqq} 0. \tag{14'}$$

(13) und (13′) besagen:

Unter allen konvexen Kurven gleicher Länge liefert das gleichseitige Dreieck und nur dieses den kleinsten Wert für das minimale Trägheitsmoment und das Quadratmittel der Entfernung der Punkte der Kurve.[9]

[7] Wir können also statt (6′) auch schreiben

$$\Delta \geqq \frac{1}{2}\pi R^2, \tag{6*}$$

die rechte Seite ist gleich dem halben Flächeninhalt eines Kreises vom Radius R. — (6*) bleibt erst recht gültig, wenn R durch das arithmetische Mittel $\frac{1}{L}\int\limits_{\mathfrak{C}} |\mathbf{d}|\,ds$ ersetzt wird.

[8] Wir betrachten es als zum Begriff der konvexen Kurve gehörig, daß diese so orientiert ist, daß bei einer Durchlaufung das Innere zur linken bleibt.

[9] Zum Beweis siehe [7], Satz 1. — Bei Beschränkung auf zentrisch-symmetrische konvexe Vergleichskurven gilt sogar

$$48I - L^3 \geqq 0,$$

wo das Gleichheitszeichen genau dann steht, wenn $\mathfrak{C}$ ein Parallelogramm ist (siehe [7], Satz 2).

[10] Ungleichung (14) wird in § 2 bewiesen. — Läßt man die Voraussetzung der Konvexität fallen, so ist (14) nicht immer richtig, da die aus (14) gefolgerte Ungleichung (16) für geeignete nichtkonvexe ebene Kurven nicht gilt, siehe Fußnote [2].

Wir wollen aus (14) einige Folgerungen ziehen. Anstelle von (14) schreiben wir auch

$$L(L^3-8\pi LF+4\pi^2 I) \underset{\mathfrak{K}}{\geqq} L(L^3-4\pi^2 I)+(L^2-4\pi F)^2; \tag{14a}$$

wegen (2) schließen wir daraus (4) (in verschärfter Form), (4) ist also für konvexe Vergleichskurven eine Folge von (14) und (2).

Schreiben wir die linke Seite von (14a) in der Form

$$2L^2(L^2-4\pi F)-L(L^3-4\pi^2 I),$$

so entnehmen wir aus (14a)

$$2L^2(L^2-4\pi F) \underset{\mathfrak{K}}{\geqq} 2L(L^3-4\pi^2 I)+(L^2-4\pi F)^2, \tag{14b}$$

eine Verschärfung der isoperimetrischen Ungleichung (5).[11] Addiert man in (14b) auf beiden Seiten $-(L^2-4\pi F)^2$, so erhält man

$$(L^2+4\pi F)(L^2-4\pi F) \underset{\mathfrak{K}}{\geqq} 2L(L^3-4\pi^2 I). \tag{14c}$$

Eine weitere Umformung von (14) ist

$$4\pi^2(LI-4F^2) \underset{\mathfrak{K}}{\geqq} L(L^3-4\pi^2 I), \tag{14d}$$

in Verbindung mit (2) folgt daraus

$$LI-4F^2 \underset{\mathfrak{K}}{\geqq} 0;\text{[11a]} \tag{15}$$

hieraus folgt mit (10)

$$\left\{\begin{aligned} &L^2Q^2-8F^2 \underset{\mathfrak{K}}{\geqq} 0,\\ &LQ-\sqrt{8}F \underset{\mathfrak{K}}{\geqq} 0.\text{[11a]}\end{aligned}\right. \tag{15'}$$

Eine andere Umformung von (14) ist

$$8\pi L(\pi I-LF) \underset{\mathfrak{K}}{\geqq} (L^2-4\pi F)^2;\text{[12]} \tag{14e}$$

[11] Aus (14) und (2) folgt also zunächst

$$2L^2(L^2-4\pi F) \underset{\mathfrak{K}}{\geqq} (L^2-4\pi F)^2$$

und daraus $L^2-4\pi F \underset{\mathfrak{K}}{\geqq} 0$, ein bemerkenswerter Schluß: Eine Größe kann deshalb nicht negativ sein, weil sie durch ihr eigenes Quadrat abgeschätzt wird!

[11a] (15) und (15′) gelten für beliebige ebene Kurven, siehe (10a) (man beachte auch Fußnote [5a]).

[12] In Verbindung mit (7) ergibt sich

$$2L^2(L^2-4\pi F) \underset{\mathfrak{K}}{\geqq} 8\pi L(\pi I-LF) \underset{\mathfrak{K}}{\geqq} (L^2-4\pi F)^2.$$

das ist eine Verschärfung der Ungleichung

(16) $$\pi I - LF \underset{\mathfrak{K}}{\geqq} 0,$$

gleichwertig mit

(16′) $$\pi Q^2 - 2F \underset{\mathfrak{K}}{\geqq} 0.$$

Aus (16), (5) und (12) gewinnen wir

$$\pi^2 I^2 \underset{\mathfrak{K}}{\geqq} L^2 F^2 \underset{\mathfrak{K},\,\mathfrak{N}}{\geqq} 4\pi F \cdot F^2,$$

folglich

(17) $$\pi I^2 - 4F^3 \underset{\mathfrak{K}}{\geqq} 0.$$

(16′) und (17) fassen wir in folgendem Satz zusammen:

Unter allen ebenen konvexen Figuren gleichen Flächeninhalts liefert der Kreis und nur dieser den kleinsten Wert für das Quadratmittel der Entfernung der Punkte der Randkurve sowie für das minimale Trägheitsmoment der Randkurve.

Aus (16) folgt im Falle $L = 2\pi$

(16″) $$I \underset{\mathfrak{K}}{\geqq} 2F \qquad (L = 2\pi),$$

und das ist nichts anderes als der eingangs genannte Satz (Ungleichung (1)).

Es sei noch hingewiesen auf einige Ungleichungen, in denen das Krümmungsträgheitsmoment J vorkommt, insbesondere kann die Ungleichung (1) in interessanter Weise erweitert werden.

ϑ bezeichne den Stützwinkel, J sei das minimale Krümmungsträgheitsmoment von $\mathfrak{C}$:

$$J = \int\limits_{\substack{\mathfrak{C}\\ O = S^*}} r^2 d\vartheta,$$

wo S^* den Steinerschen Krümmungsschwerpunkt bezeichnet.[13] Dann gilt

(18) $$2\pi F + \pi J - L^2 \underset{\mathfrak{K}}{\geqq} 0,^{14}$$

oder anders geschrieben:

(18′) $$2\pi J - L^2 \underset{\mathfrak{K}}{\geqq} L^2 - 4\pi F;$$

daraus folgt wegen (5)

(19) $$2\pi J - L^2 \underset{\mathfrak{K}}{\geqq} 0.^{15}$$

[13] Vergleiche [8], S. 347, sowie das Literaturverzeichnis von [8].
[14] Beweis in § 2.
[15] Siehe [8], Satz 5.

Weiter gilt

(20) $$\pi L^2 - 8J \underset{\Re}{\geqq} 0.$$ [15]

Aus (19) und (2) folgen (in Verbindung mit (10))

(21) $$LJ - 2\pi I \underset{\Re}{\geqq} 0,$$ [16]

(21′) $$J - \pi Q^2 \underset{\Re}{\geqq} 0,$$

(22) $$J^3 - 2\pi I^2 \underset{\Re}{\geqq} 0.$$

Mit $L = 2\pi$ geht (21) über in

$$J \underset{\Re}{\geqq} I \qquad (L = 2\pi),$$

in Verbindung mit (16″) ergibt sich

(23) $$\int\limits_{\substack{\mathfrak{C} \\ O=S^*}} r^2 d\vartheta \underset{\Re}{\geqq} \int\limits_{\substack{\mathfrak{C} \\ O=S}} r^2 ds \underset{\Re}{\geqq} \int\limits_{\substack{\mathfrak{C} \\ O \text{ beliebig}}} r^2 d\varphi$$

$$\left(\int d\vartheta = \int ds = \int d\varphi = 2\pi\right).$$

Aus (2) gewinnt man durch Multiplikation mit $\dfrac{Q^4}{2L}$

$$2\left(\frac{LQ^2}{2}\right)^2 - \pi^2 \frac{2IQ^4}{L} \underset{\Re}{\geqq} 0,$$

oder, wegen (10),

$$2I^2 - \pi^2 Q^6 \underset{\Re}{\geqq} 0,$$

folglich

(24) $$\sqrt{2}I - \pi Q^3 \underset{\Re}{\geqq} 0.$$ [17]

Ebenso schließt man aus (13)

(25) $$3\sqrt{3}Q^3 - 2I \underset{\mathfrak{D}}{\geqq} 0.$$

Ungleichungen (24) und (25) besagen:

Unter allen konvexen Kurven gleichen minimalen Trägheitsmomentes liefert der Kreis und nur dieser den größten Wert und das gleichseitige Dreieck und nur dieses den kleinsten Wert für das Quadratmittel der Entfernung der Punkte der Kurve.

Der erste Teil des Satzes gilt auch für Raumkurven.

Schließlich kann man diejenigen der gewonnenen Ungleichungen, in

[16] Vergleiche [7], Satz 6.

[17] Ungleichung (24) gilt auch für Raumkurven.

denen nur zwei der betrachteten Größen vorkommen und in denen im Falle des Kreises Gleichheit eintritt (also die Ungleichungen (2), (5), (2'), (16'), (17), (19), (21'), (22), (24)) unter Benutzung einer von PÓLYA und SZEGŐ[18] eingeführten zweckmäßigen Schreibweise übersichtlich zusammenfassen:

(26) $$\bar{J} \underset{\mathfrak{K}}{\geqq} \bar{L} \underset{\mathfrak{K}}{\geqq} \bar{I} \underset{\mathfrak{K}}{\geqq} \bar{Q} \underset{\mathfrak{K}}{\geqq} \bar{F}.$$

*

Unter allen angegebenen Ungleichungen zwischen L, F und I für konvexe Kurven sind unabhängig nur die folgenden vier:

(2) $$L^3 - 4\pi^2 I \underset{\mathfrak{K}}{\geqq} 0,$$

(12) $$F \underset{\mathfrak{K}}{\geqq} 0,$$

(13) $$54I - L^3 \underset{\mathfrak{D}}{\geqq} 0,$$

(14) $$8\pi^2 LI - L^4 - (4\pi F)^2 \underset{\mathfrak{K}}{\geqq} 0;$$

(2) gilt allgemein für Raumkurven, (12), (13), (14) aber sind ohne Voraussetzung der Konvexität nicht immer erfüllt. Die vier genannten Ungleichungen bilden jedoch kein vollständiges System, das heißt: Wenn die drei Zahlen $L > 0, F, I$ alle vier Ungleichungen erfüllen, so folgt daraus *nicht* die Existenz einer konvexen Kurve, deren Länge mit L, deren Flächeninhalt mit F und deren minimales Trägheitsmoment mit I übereinstimmt: Man setze etwa $L = 2\pi$, $I = 2\pi$, $0 \leqq F < \pi$, dann sind alle vier Ungleichungen erfüllt, wegen (2) müßte der Kreis vorliegen, dann müßte aber $F = \pi$ sein. Aus Kompaktheitsgründen[19] gibt es auch Zahlentripel $L > 0$, F, I, welche nicht zu einer konvexen Kurve gehören und alle vier Ungleichungen mit dem $>$-Zeichen erfüllen.

§ 2. Beweise

BEWEIS VON UNGLEICHUNG (14):

$$8\pi^2 LI - L^4 - (4\pi F)^2 \underset{\mathfrak{K}}{\geqq} 0.$$

Wir führen Stützkoordinaten ϑ, p ein (Fig. 3) und zeigen zunächst: Für eine beliebige konvexe Kurve $\mathfrak{C}$ gilt

(27) $$I \equiv \int_{\mathfrak{C}} r^2 ds = \frac{1}{2} \left\{ \int_{\mathfrak{C}} p^2 ds + \int_{\mathfrak{C}} p^3 d\vartheta \right\}.$$

[18] [4], 1.5 (S. 4), vergleiche auch 1.15 (S. 10); siehe auch [9], S. 358.
[19] Man benutze den Auswahlsatz von BLASCHKE, [1], S. 62.

Zum Beweise nehmen wir der Einfachheit halber an, daß $p(\vartheta)$ zweimal stetig differenzierbar sei; das ist keine Einschränkung, weil jede konvexe Kurve durch Kurven mit der genannten Eigenschaft beliebig genau approximiert werden kann, und dann folgt (27) durch Grenzübergang. — Wir beachten:

$$r^2 = p^2 + p'^2, \quad ds = (p + p'')d\vartheta,$$
$$\int_{\mathfrak{C}} p\,d\vartheta = L, \quad \int_{\mathfrak{C}} p\,ds = 2F.$$

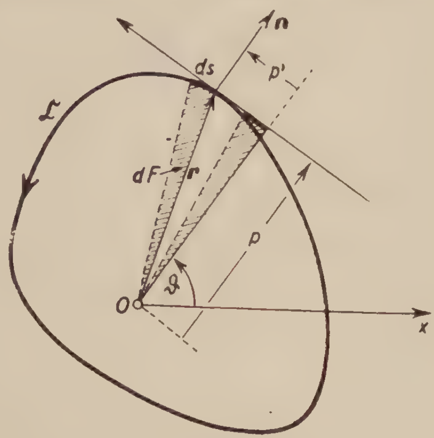


Fig. 3

Es ist $I = \int_{\mathfrak{C}} (p^2 + p'^2)ds$. Wir formen das zweite Glied $\int_{\mathfrak{C}} p'^2 ds$ um:

$$(28)\qquad \int_{\mathfrak{C}} p'^2 ds = \int_{\mathfrak{C}} p'^2(p + p'')d\vartheta = \int_{\mathfrak{C}} p'^2 p\,d\vartheta,$$

da $\int_{\mathfrak{C}} p'^2 p''\,d\vartheta = \left[\frac{1}{3} p'^3\right]_{\vartheta=0}^{2\pi} = 0$ ist. Partielle Integration liefert

$$\int_{\mathfrak{C}} p'^2 ds = \int_{\mathfrak{C}} pp' \cdot p'd\vartheta = [pp' \cdot p]_{\vartheta=0}^{2\pi} - \int_{\mathfrak{C}} \frac{d}{d\vartheta}(pp') \cdot p\,d\vartheta,$$

wo wieder $[pp' \cdot p]_{\vartheta=0}^{2\pi} = 0$ ist, so daß

$$(29)\qquad \int_{\mathfrak{C}} p'^2 ds = -\int_{\mathfrak{C}} p'^2 p\,d\vartheta - \int_{\mathfrak{C}} p^2 p''\,d\vartheta.$$

Aus (28) und (29) ergibt sich durch Addition

$$(30)\qquad \int_{\mathfrak{C}} p'^2 ds = -\frac{1}{2}\int_{\mathfrak{C}} p^2 p''\,d\vartheta = -\frac{1}{2}\left\{\int_{\mathfrak{C}} p^2(p + p'')d\vartheta - \int_{\mathfrak{C}} p^3 d\vartheta\right\} =$$
$$= -\frac{1}{2}\left\{\int_{\mathfrak{C}} p^2 ds - \int_{\mathfrak{C}} p^3 d\vartheta\right\}.$$

Daraus ergibt sich

$$I = \int_{\mathfrak{C}} p^2 ds + \int_{\mathfrak{C}} p'^2 ds = \frac{1}{2}\left\{\int_{\mathfrak{C}} p^2 ds + \int_{\mathfrak{C}} p^3 d\vartheta\right\},$$

das ist aber gerade die zu beweisende Relation (27).

Nun wenden wir mehrmals die Schwarzsche Ungleichung an:

$$L \int_{\mathfrak{C}} p^2 ds = \int_{\mathfrak{C}} ds \int_{\mathfrak{C}} p^2 ds \geqq \left\{ \int_{\mathfrak{C}} p ds \right\}^2 = 4F^2,$$

$$L \int_{\mathfrak{C}} p^3 d\vartheta = \int_{\mathfrak{C}} p d\vartheta \int_{\mathfrak{C}} p^3 d\vartheta \geqq \left\{ \int_{\mathfrak{C}} p^2 d\vartheta \right\}^2,$$

$$2\pi \int_{\mathfrak{C}} p^2 d\vartheta = \int_{\mathfrak{C}} d\vartheta \int_{\mathfrak{C}} p^2 d\vartheta \geqq \left\{ \int_{\mathfrak{C}} p d\vartheta \right\}^2 = L^2;$$

daraus ergibt sich

$$\int_{\mathfrak{C}} p^2 ds \geqq \frac{4F^2}{L},$$

$$\int_{\mathfrak{C}} p^3 d\vartheta \geqq \frac{1}{L}\left(\frac{L^2}{2\pi}\right)^2 = \frac{L^3}{4\pi^2},$$ [20]

wo in jedem Falle das Gleichheitszeichen genau dann steht, wenn p konstant ist, also genau dann, wenn $\mathfrak{C}$ ein Kreis (mit dem Ursprung als Mittelpunkt) ist. — Durch Addition ergibt sich

$$I \underset{\mathfrak{K}}{\geqq} \frac{1}{2}\left\{\frac{4F^2}{L} + \frac{L^3}{4\pi^2}\right\}$$

oder

$$8\pi^2 L I - L^4 - (4\pi F)^2 \underset{\mathfrak{K}}{\geqq} 0,$$

wie zu beweisen war.

*

Beweis von Ungleichung (18):

$$2\pi F + \pi J - L^2 \underset{\mathfrak{K}}{\geqq} 0.$$

Wir benutzen Stützkoordinaten mit dem Krümmungsschwerpunkt als Ursprung. — Es ist

$$2F = \int_{\mathfrak{C}} p ds = \int_{\mathfrak{C}} p(p+p'') d\vartheta = \int_{\mathfrak{C}} (p^2 - p'^2) d\vartheta,$$

$$J = \int_{\mathfrak{C}} r^2 d\vartheta = \int_{\mathfrak{C}} (p^2 + p'^2) d\vartheta,$$

[20] Dasselbe ergibt sich bei Anwendung der Ungleichung für Potenzmittel:

$$\left\{\frac{1}{2\pi}\int p^3 d\vartheta\right\}^{\frac{1}{3}} \geqq \frac{1}{2\pi}\int p d\vartheta = \frac{L}{2\pi}.$$

und durch Addition ergibt sich

$$2F+J=2\int_{\mathfrak{C}} p^2 d\vartheta;$$

durch Anwendung der Schwarzschen Ungleichung gewinnt man

$$2F+J \geqq 2\,\frac{1}{2\pi}\left\{\int_{\mathfrak{C}} p\,d\vartheta\right\}^2 = \frac{1}{\pi}L^2, \tag{31}$$

wo das Gleichheitszeichen nur im Falle $p=\text{const}$, also nur im Falle des Kreises steht. — Aus (31) folgt sofort die Behauptung.

(Eingegangen am 8. September 1959.)

Literaturverzeichnis

[1] W. Blaschke, *Kreis und Kugel* (Leipzig, 1916).
[2] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities* (Cambridge, 1934).
[3] A. Hurwitz, Sur le problème des isopérimètres, *Comptes Rendus Acad. Sci. Paris*, **132** (1901), S. 401—402.
[4] G. Pólya and G. Szegő, *Isoperimetric inequalities in mathematical physics* (Princeton, 1951).
[5] L. Rédei und B. Sz.-Nagy, Eine Verallgemeinerung der Inhaltsformel von Heron, *Publ. Math. Debrecen*, **1** (1949), S. 42—50.
[6]—[8] H. Sachs, Über eine Klasse isoperimetrischer Probleme, 1.—3. Mitteilung, *Wiss. Z. Univ. Halle*, Math.-Nat Reihe, VIII/1 (1958), S. 121—126, 127—133 und VIII/3 (1959), S. 345—350.
[9] H. Sachs, Zur Theorie gewisser geometrischer Funktionale und zugehöriger isoperimetrischer Probleme, *Wiss. Z. Univ. Halle*, Math.-Nat. Reihe, VIII/3 (1959), S. 357—364.

Ferner sei allgemein hingewiesen auf die folgenden mit umfassenden Literaturverzeichnissen versehenen Bücher:

L. Fejes Tóth, *Lagerungen in der Ebene, auf der Kugel und im Raum* (Berlin—Göttingen—Heidelberg, 1953).
H. Hadwiger, *Altes und Neues über konvexe Körper* (Basel und Stuttgart, 1955).
H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie* (Berlin—Göttingen—Heidelberg, 1957).

NOTES ON ABELIAN GROUPS. II

By

L. FUCHS (Budapest)

(*Presented by* L. RÉDEI)

§ 5. p-basic subgroups of arbitrary abelian groups

KULIKOV [8] introduced the notion of basic subgroups of abelian p-groups which has proved to be one of the most important notions in the theory of p-groups of arbitrary power. Basic subgroups can be defined in any module over the ring of p-adic integers, or, more generally, over any discrete valuation ring. Here we want to give a generalization of basic subgroups to any group so that it coincides with the old concept whenever the group is primary. In the general case, to every prime p, one can define p-basic subgroups where in the definition the prime p plays a distinguished role. The p-basic subgroups are not isomorphic for different primes, but are uniquely determined (up to isomorphism) by the group and the prime p. We shall see that p-basic subgroups are useful in certain investigations.

Let G be an arbitrary (abelian) group[1] and p an arbitrary, but fixed prime. We call a subset $[x_\lambda]_{\lambda\in\Lambda}$ of G, not containing 0, *p-independent*, if for any finite subset $x_1,\dots,x_k$ a relation

$$n_1x_1+\dots+n_kx_k\in p^rG \qquad (n_ix_i\neq 0)$$

implies

$$p^r|n_i \qquad (i=1,2,\dots,k).$$

LEMMA 1. *A p-independent set is independent.*

Assume $[x_\lambda]$ is p-independent and $n_1x_1+\dots+n_kx_k=0$ where $n_ix_i\neq 0$. Then $n_1x_1+\dots+n_kx_k\in p^rG$ for every r whence each n_i is divisible by every power of p. Thus $n_i=0$ and the arising contradiction establishes the statement.

A subgroup H of G is called *p-pure*[2] if

$$p^rH=H\cap p^rG \quad \text{for} \quad r=1,2,\dots,$$

[1] For the basic facts and terminology we refer to our book [3]. — Let us note that by $\mathfrak{C}(n)$ we denote the cyclic group of order $n=1,2,\dots$ or ∞, by $\mathfrak{C}(p^\infty)$ the quasicyclic group, by $\mathfrak{C}$ the group of all rotations of finite order of the circle, by $\mathfrak{R}$ the group of rationals and by $\mathfrak{J}$ the group of p-adic integers.

[2] Remember that a subgroup H of G is called pure if $nH=H\cap nG$ for every integer n. This is equivalent to the fact that $p^rH=H\cap p^rG$ for every prime p and for every positive integer r.

or, otherwise expressed, the solvability of an equation $p^r x = h (\in H)$ in G implies its solvability in H.

LEMMA 2. *The subgroup H generated by a p-independent subset $[x_\lambda]$ of G is p-pure in G. Conversely, if an independent set containing but elements of p-power order and or of order ∞ generates a p-pure subgroup, then it is p-independent.*

Assume that $h \in H \cap p^r G$ where H is the subgroup generated by a p-independent set $[x_\lambda]$ of G. Then $h = n_1 x_1 + \cdots + n_k x_k \in p^r G$ where we may assume that $n_i x_i \neq 0$. By p-independence there exist integers n_i' with $n_i = p^r n_i'$ and so $h = p^r(n_1' x_1 + \cdots + n_k' x_k) \in p^r H$, i. e. $p^r H = H \cap p^r G$.

Next let $[x_\lambda]$ be an independent set such that the x_λ are of p-power order or of order ∞ and $H = \{\ldots, x_\lambda, \ldots\}$ is p-pure in G. Assume $n_1 x_1 + \cdots + n_k x_k \in p^r G$ $(n_i x_i \neq 0)$. By p-purity we have $n_1 x_1 + \cdots + n_k x_k = p^r(n_1' x_1 + \cdots + n_k' x_k)$ whence by independence $n_i x_i = p^r n_i' x_i$. From the hypothesis on the orders of x_λ we obtain $p^r | n_i$. Q. e. d.

Calling a group G *p-divisible*[3] if it satisfies $pG = G$, we define: a subgroup B of an arbitrary group G will be called a *p-basic subgroup of G* if

(i) B is the direct sum of cyclic p-groups and/or infinite cyclic groups, $B = \Sigma\{a_\lambda\}$;

(ii) B is a p-pure subgroup of G;

(iii) the factor group G/B is p-divisible.

A basis $[a_\lambda]$ of B is said to be a p-basis of G.

LEMMA 3. *$[a_\lambda]$ is a p-basis of G if and only if it is a maximal p-independent set containing but elements of p-power order and/or infinite order.*

Assume that $[a_\lambda]$ is a maximal p-independent set of the indicated property. Then (i) and (ii) are obvious. In order to verify (iii), let $g \neq 0$ be an arbitrary element of G; by maximality there exists a relation $n_0 g + n_1 a_1 + \cdots + n_k a_k \in p^r G$ for some r such that $n_0 g \neq 0$, $n_i a_i \neq 0$ for $i = 1, \ldots, k$ and $p^r \nmid n_j$ for some j. By the p-independence of the a_λ surely $p^r \nmid n_0$. There is no loss of generality in assuming that $n_0 = p^s$ with $s < r$. Now, $n_1 a_1 + \cdots + n_k a_k \in p^s G$ implies that $n_i = p^s n_i'$ for every i. Thus $p^s(g + n_1' a_1 + \cdots + n_k' a_k) = p^r g'$ for some $g' \in G$, i. e. $p^s(g - p^{r-s} g' + n_1' a_1 + \cdots + n_k' a_k) = 0$. Suppose that we have proved the divisibility of the elements of G by p mod B provided their orders are powers of p and less than p^r. Then it follows that $g - p^{r-s} g' + n_1' a_1 + \cdots + n_k' a_k$ and hence g is divisible by p mod B. Consequently, divisibility by p mod B holds also for elements of order p^r, and so, by induction, for every element of p-power order. If g is of infi-

[3] Cf. e. g. SPECHT [10].

nite order, then — knowing the divisibility of elements of finite order by p — our inference leads to the conclusion that it is divisible by p mod B, thus $[a_\lambda]$ is a p-basis of G.

Conversely, if $[a_\lambda]$ is a p-basis, then by Lemma 2 it is a p-independent set. From (iii) we infer that to any non-zero $g \in G$ there exists a relation of the form $g + n_1 a_1 + \cdots + n_k a_k \in pG$, thus if we enlarge the set $[a_\lambda]$ by g, we obtain no longer a p-independent set.

The main result on p-basic subgroups is

THEOREM 5. 1. *Every group G contains p-basic subgroups for every prime p. The p-basic subgroups of G (for the same prime p) are all isomorphic.*

p-independence being a property of finite character, there exist maximal p-independent sets containing but elements of p-power and/or infinite order. By Lemma 3, we are ready with the proof of the existence.

In order to prove unicity, we begin with showing that $G = \{B, p^n G\}$ for every p-basic subgroup B of G and for every natural integer n. In fact, every element of G is divisible by p^n mod B. Hence and because of the p-purity of B we obtain

$$B/p^n B = B/(B \cap p^n G) \sim \{B, p^n G\}/p^n G = G/p^n G.$$

The number of components of order p^r in B is the same as the corresponding number in $B/p^n B$ for $n > r$ and hence in $G/p^n G$. — Next we verify the invariance of the number of infinite cyclic components in p-basic subgroups of G. We first prove that $pB_0 = B_0 \cap \{T, pG\}$ where B_0 denotes the direct sum of the infinite cyclic components of B and T is the maximal p-subgroup of G. The inclusion $\subseteq$ being trivial, let $b_0 \in B_0 \cap \{T, pG\}$ and we have to show that $b_0 \in pB_0$. Write $b_0 = d + pg$ $(d \in T,\ g \in G)$; for some n we have $p^n b_0 = p^{n+1} g \in B_0 \cap p^{n+1} G = p^{n+1} B_0$, i. e. $p^n b_0 = p^{n+1} b$ for some $b \in B_0$. Hence $b_0 = pb \in pB_0$, in accordance with the torsion freeness of B_0. We infer

$$B_0/pB_0 = B_0/(B_0 \cap \{T, pG\}) \cong \{B_0, T, pG\}/\{T, pG\} = G/\{T, pG\}$$

considering that $\{B_0, T, pG\} \supseteq \{B, pG\} = G$. Thus the rank of B_0 is independent of the special choice of B. This completes the proof.

Let us note that it is easy to show that in a p-basis of G the elements of finite order form a p-basis of T^* and the elements of infinite order form a p-basis of G/T^* where T^* is the maximal torsion subgroup of G.

We shall make use of the following simple fact:

LEMMA 4. *Let H be a p-pure subgroup of G, $[x_\lambda]$ a p-basis of H and $[y^*_\mu]$ a p-basis of $G^* = G/H$. If $y_\mu \in G$ is an element in the coset y^*_μ having the same order as y^*_μ, then $[x_\lambda, y_\mu]$ is a p-basis of G.*

Since the orders of the y^*_μ are powers of p or infinite, such elements y_μ exist by p-purity.[4] The set $[x_\lambda, y_\mu]$ is p-independent for if $n_1x_1 + \cdots + n_kx_k + m_1y_1 + \cdots + m_ty_t = p^rg$ $(g \in G)$, then passing mod H we obtain $m_1y^*_1 + \cdots + m_ty^*_t = p^rg^*$ whence, for every j, either $m_jy^*_j = 0$ (and so $m_jy_j = 0$) or $p^r | m_j$. In any case we have a relation $n_1x_1 + \cdots + n_kx_k = p^rh$ for some $h \in G$. Hence, for every i, either $n_ix_i = 0$ or $p^r | n_i$, establishing p-independence. In order to verify maximality, note that to every $g \in G$ there exists a linear combination $m_1y_1 + \cdots + m_ty_t$ such that $g + m_1y_1 + \cdots + m_ty_t \in \{pG, H\}$, i.e. $g + m_1y_1 + \cdots + m_ty_t = pg' + h$ for some $g' \in G$, $h \in H$. To h there exists a relation $h - n_1x_1 - \cdots - n_kx_k = ph'$ with $h' \in H$, showing that $g + n_1x_1 + \cdots + n_kx_k + m_1y_1 + \cdots + m_ty_t \in pG$. Therefore, the maximality of the p-independent set $[x_\lambda, y_\mu]$ is established.

Let us turn to some applications.

THEOREM 5.2. *If T is a p-group and $B = \sum \{a_\lambda\}$ is a p-basic subgroup of G, then*

$$G \otimes T \cong B \otimes T \cong \sum_\lambda (\{a_\lambda\} \otimes T).$$

Remember that if $0 \to A \to B \to C \to 0$ is a pure exact sequence,[5] then for any group T, the induced sequence $0 \to A \otimes T \to B \otimes T \to C \otimes T \to 0$ is likewise pure exact. It is easy to prove the following analogue of this statement: if $0 \to A \to B \to C \to 0$ is a p-pure exact sequence and T is a p-group, then the sequence $0 \to A \otimes T \to B \otimes T \to C \otimes T \to 0$ is exact. Applying this fact to the p-pure exact sequence $0 \to B \to G \to G/B \to 0$, in view of $(G/B) \otimes T = 0$ we obtain that $0 \to B \otimes T \to G \otimes T \to 0$ is exact, i.e. $G \otimes T \cong B \otimes T$. The second isomorphism of the theorem follows from the permutability of forming direct sums and tensor products.

THEOREM 5.3. (HARRISON [6a].) *If H is a pure subgroup of G and T is a torsion group, then*

$$G \otimes T \cong H \otimes T + (G/H) \otimes T.$$

It suffices to verify this isomorphism only for p-groups T. For p-groups T it is a simple consequence of Lemma 4 and the preceding theorem.

[4] For the corresponding fact on purity see [7] or [3].

[5] That is to say, A is a pure subgroup of B. If A is p-pure in B, we call the sequence p-pure exact. For the cited property see [5].

THEOREM 5.4. *If B is a p-basic subgroup of G, then*

$$\mathrm{Hom}(G, \mathcal{C}(p^\infty)) \simeq \mathrm{Hom}(G/B, \mathcal{C}(p^\infty)) + \mathrm{Hom}(B, \mathcal{C}(p^\infty)).$$

The maximal p-subgroup T of G/B is clearly a divisible p-group, hence a direct summand of G/B. Any homomorphism of G/B into $\mathcal{C}(p^\infty)$ maps q-subgroups (q a prime, $q \neq p$) upon 0, consequently, $\mathrm{Hom}(G/B, \mathcal{C}(p^\infty)) \simeq \mathrm{Hom}(T, \mathcal{C}(p^\infty)) + \mathrm{Hom}(J, \mathcal{C}(p^\infty))$ where J is the factor group of G/B with respect to its maximal torsion subgroup. Since J is p-divisible and in $\mathcal{C}(p^\infty)$ division by q is possible and unique, $\mathrm{Hom}(J, \mathcal{C}(p^\infty))$ is divisible. Further, $\mathrm{Hom}(T, \mathcal{C}(p^\infty)) \simeq \mathrm{Hom}(\sum \mathcal{C}(p^\infty), \mathcal{C}(p^\infty)) \simeq \simeq \sum^* \mathrm{Hom}(\mathcal{C}(p^\infty), \mathcal{C}(p^\infty)) \simeq \sum^* \mathfrak{S}$. Thus $\mathrm{Hom}(G/B, \mathcal{C}(p^\infty))$ is algebraically compact, i. e. it has the property that it is a direct summand of any group containing it as a p-pure subgroup.[6]

We know that the p-pure exact sequence $0 \to B \to G \to G/B \to 0$ implies that, for any group V, the sequence $0 \to \mathrm{Hom}(G/B, V) \to \mathrm{Hom}(G, V) \to \to \mathrm{Hom}(B, V)$ likewise p-pure exact.[7] If V is a divisible group, then we can complete this by $\to 0$. By what has been said in the preceding paragraph it follows that $\mathrm{Hom}(G/B, \mathcal{C}(p^\infty))$ is a direct summand of $\mathrm{Hom}(G, \mathcal{C}(p^\infty))$ whence the assertion.

THEOREM 5.5. *If H is a p-pure subgroup of G, then*

$$\mathrm{Hom}(G, \mathcal{C}(p^\infty)) \simeq \mathrm{Hom}(H, \mathcal{C}(p^\infty)) + \mathrm{Hom}(G/H, \mathcal{C}(p^\infty)).$$

Proof again by making use of algebraic compactness of $\mathrm{Hom}(G/H, \mathcal{C}(p^\infty))$.

§ 6. Groups whose extensions by torsion free groups split

J. Łoś [9] discussed groups which are direct summands of groups containing them as pure subgroups and BALCERZYK [1] proved that these coincide with the algebraically compact groups defined formerly by KAPLANSKY [7]. Here we intend to deal with a generalization of algebraically compact groups, namely with groups which are direct summands of groups containing them such that the corresponding factor groups are torsion free. The importance of such groups is motivated by the fact the group $\mathrm{Ext}(L, K)$ of extensions of K by L has this property for any groups K and L.

Call a group G an *A-group* (algebraically compact) if G is a direct summand of any group H whenever H contains G as a pure subgroup.

[6] Cp. Łoś [9]. — For the algebraic compactness of homomorphism groups into $\mathcal{C}(p^\infty)$ see [4].

[7] See [5].

G will be said to be a *B-group*[7a] if G is a direct summand of H if H contains G such that the factor group H/G is torsion free, i. e. every extension of G by a torsion free group is splitting. By definition, every A-group is at the same time a B-group, but — as we shall see below — not conversely.

The following simple properties of B-groups may be mentioned:

(a) *G is a B-group if* (and clearly only if) *it has the property that every extension of G by $\mathfrak{R}$ is splitting.*

Every torsion free group J can be embedded in a direct sum of groups $\mathfrak{R}$, thus $0 \to J \to \sum \mathfrak{R}$ is an exact sequence. This implies that[8] $\operatorname{Ext}(\sum \mathfrak{R}, G) \to \operatorname{Ext}(J, G) \to 0$ is exact for every group G. If G has the stated property, then $\operatorname{Ext}(\sum \mathfrak{R}, G) = \sum^* \operatorname{Ext}(\mathfrak{R}, G) = 0$ and therefore $\operatorname{Ext}(J, G) = 0$ for every torsion free J, i. e. G is a B-group.

(b) *Every homomorphic image of a B-group is again a B-group.*

If $G \to H \to 0$ is an exact sequence, then so is $\operatorname{Ext}(J, G) \to \operatorname{Ext}(J, H) \to 0$, too. If in the last sequence the first group vanishes for every torsion free J, then so does the second.

(c) *If G is a B-group, and H is a subgroup of G such that the factor group G/H is reduced, then H is a B-group.*

The exact sequence $0 \to H \to G \to G/H \to 0$ implies the exactness of $\cdots \to \operatorname{Hom}(\mathfrak{R}, G/H) \to \operatorname{Ext}(\mathfrak{R}, H) \to \operatorname{Ext}(\mathfrak{R}, G) \to \cdots$. By hypothesis, the first and the third groups are here 0, therefore the second is also 0. (a) implies the assertion.

(d) *If H and G/H are both B-groups, then G is a B-group.*

From the exact sequence $0 \to H \to G \to G/H \to 0$ we obtain the exact sequence $\operatorname{Ext}(\mathfrak{R}, H) \to \operatorname{Ext}(\mathfrak{R}, G) \to \operatorname{Ext}(\mathfrak{R}, G/H)$. By hypothesis, the first and the third groups vanish, therefore $\operatorname{Ext}(\mathfrak{R}, G) = 0$, as stated.

(e) *The complete direct sum of groups G_λ ($\lambda \in \Lambda$) is a B-group if and only if every component G_λ is a B-group.*

This is obvious in view of the isomorphism $\operatorname{Ext}(\mathfrak{R}, \sum^* G_\lambda) \simeq \sum^* \operatorname{Ext}(\mathfrak{R}, G_\lambda)$.

In order to investigate some deeper properties of B-groups, we recall the definition of Ulm subgroups and Ulm factors of a group.[9] The elements of a group G which are divisible by every natural integer n, form a subgroup G^1 of G; thus $G^1 = \bigcap_n nG$. We call G^1 the first Ulm subgroup of G.

[7a] Only after the manuscript had been finished did I notice that reduced B-groups (under the name "co-torsion" groups) have been discussed by D. K. Harrison [6]. His and my results have not much interrelation.

[8] For the needed facts on Homological Algebra we refer to [2].

[9] This general definition may be found in Specht [10].

If G^β for every $\beta<\alpha$ are already defined, then — in case α is an isolated ordinal — let G^α be the first Ulm subgroup of $G^{\alpha-1}$ and — in case α is a limit ordinal — let G^α be the intersection of all G^β with $\beta<\alpha$; G^α is called the *αth Ulm subgroup of G*. The factor groups $G_\alpha=G^\alpha/G^{\alpha+1}$ are said to be the *Ulm factors of G*. The first Ulm subgroup of any factor G_α is zero.

(f) *The Ulm subgroups of a B-group are again B-groups.*

The factor groups G/G^α are easily seen to be reduced groups, consequently, (f) is an immediate consequence of (c).

(g) *A reduced B-group G is an A-group if and only if its first Ulm subgroup G^1 is zero.*

From the exact sequence $0\to\mathfrak{C}(\infty)\to\mathfrak{R}\to\mathfrak{C}\to 0$ we obtain the exact sequence $\cdots\to \mathrm{Hom}(\mathfrak{R},G)\to \mathrm{Hom}(\mathfrak{C}(\infty),G)\to \mathrm{Ext}(\mathfrak{C},G)\to \mathrm{Ext}(\mathfrak{R},G)\to\cdots$. Here the first and the last groups are 0 by assumption, while the second is isomorphic to G. Thus we obtain $\mathrm{Ext}(\mathfrak{C},G)\simeq G$ for any B-group G. Remember that a group G is an A-group if and only if $\mathrm{Ext}(\mathfrak{R},G)=0$ and the irst Ulm subgroup of $\mathrm{Ext}(\mathfrak{C},G)$ vanishes;[10] thus (g) holds, in fact.

(h) *The Ulm factors of a B-group are A-groups.*

This is evident from (f), (b) and (g).

It is natural to raise the question: is a group necessarily a B-group if its Ulm factors are A-groups? If the group has but a finite number of Ulm factors, then the answer is trivially yes. But if the number of Ulm factors is infinite, then a counterexample shows that in general the answer is no. The idea of constructing such a counterexample is the same as the idea of proof of Lemma 38. 1 in my book [3]; since it is rather complicated, we omit the details.

(i) *A torsion group G is a B-group if and only if it is a direct sum of a divisible torsion group and a bounded group (thus it is an A-group).*

The first Ulm factor G_1 of G is a torsion A-group without elements of infinite height, hence it is a bounded group.[11] But if the first Ulm factor of a group G is bounded, then G is the direct sum of a bounded group and a divisible group whence the statement.

(j) *A torsion free group is a B-group if and only if it is an A-group.*

A torsion free group G satisfies $G^1=G^2$ because of unicity of division by integers. Consequently, G^1 is the maximal divisible subgroup of G, and therefore $G\simeq G_1+G^1$ is the direct sum of two A-groups whenever it is a B-group.

[10] See [3].

[11] Cf. Kaplansky [7].

As an application we verify (cp. HARRISON [6]):

THEOREM. *The group* Ext (L, K) *of extensions of* K *by* L *is a B-group for arbitrary* K *and* L.

We make use of the following isomorphism:[12]

$$\mathrm{Ext}(A, \mathrm{Ext}(B, C)) \simeq \mathrm{Ext}(\mathrm{Tor}(A, B), C).$$

Apply this for $A = \mathfrak{R}$, $B = L$, $C = K$, and observe that $\mathrm{Tor}(\mathfrak{R}, K) = 0$ in view of the torsion freeness of $\mathfrak{R}$. Therefore $\mathrm{Ext}(\mathfrak{R}, \mathrm{Ext}(L, K)) = 0$, q. e. d.

Remember that the extensions of K by L which belong to the first Ulm subgroup of $G = \mathrm{Ext}(K, L)$ may be characterized by the property that in them K is a pure subgroup.[13] Thus we have:

COROLLARY. $\mathrm{Ext}(K, L)/\mathrm{Ext}(K, L; \aleph_0)$ *is algebraically compact.*

Finally, let us mention that (i) and (j) give information about groups $\mathrm{Ext}(K, L)$ which are torsion resp. torsion free.

The following example shows that the groups $\mathrm{Ext}(L, K)$ are in general no A-groups.[14] Let L be an arbitrary torsion group. It is easy to see that there exists a direct sum of cyclic p-groups T such that $L \simeq T/K$ for some pure subgroup K of T. This means that $\mathrm{Ext}(L, K; \aleph_0) \neq 0$. On the other hand, $\mathrm{Ext}(L, K)$ is reduced in the mentioned case. In fact, if we apply the isomorphism [2]

$$\mathrm{Ext}(A, \mathrm{Hom}(B, C)) + \mathrm{Hom}(A, \mathrm{Ext}(B, C)) \simeq \mathrm{Ext}(A \otimes B, C) + \mathrm{Hom}(\mathrm{Tor}(A, B), C)$$

to the groups $A = \mathfrak{R}$, $B = L$, $C = K$, and observe that now both $\mathfrak{R} \otimes L$ and $\mathrm{Tor}(\mathfrak{R}, L)$ vanish, we obtain $\mathrm{Hom}(\mathfrak{R}, \mathrm{Ext}(L, K)) = 0$. This is equivalent to the fact that $\mathrm{Ext}(L, K)$ is reduced. Consequently, in this case $\mathrm{Ext}(L, K)$ is a reduced group with non-vanishing first Ulm subgroup, i. e. it cannot be an A-group.

It is an open problem which extensions of K by L correspond to the elements of the higher Ulm subgroups of $\mathrm{Ext}(L, K)$.

(Received 17 September 1959)

[12] See CARTAN—EILENBERG [2].

[13] Thus $G^{1} = \mathrm{Ext}(L, K; \aleph_0)$; cf. [3].

[14] In [5] it has been shown that $\mathrm{Ext}(L, K)$ is an A-group whenever K is torsion free.

References

[1] S. Balcerzyk, On algebraically compact groups of I. Kaplansky, *Fundamenta Math.*, **44** (1957), pp. 91–93.

[2] H. Cartan and S. Eilenberg, *Homological algebra* (Princeton, 1956).

[3] L. Fuchs, *Abelian groups* (Budapest, 1958).

[4] L. Fuchs, On character groups of discrete abelian groups, *Acta Math. Acad. Sci. Hung.*, **10** (1959), pp. 133–140.

[5] L. Fuchs, Notes on abelian groups. I, *Annales Univ. Sci. Budapest*, Sectio Math., **2** (1959), pp. 5–23.

[6] D. K. Harrison, Infinite abelian groups and homological methods, *Annals of Math.*, **69** (1959), pp. 366–391.

[6a] D. K. Harrison, Two of the problems of L. Fuchs, *Publ. Math. Debrecen* (under press).

[7] I. Kaplansky, *Infinite abelian groups* (Ann Arbor, 1954).

[8] Л. Я. Куликов, К теории абелевых групп произвольной мощности, Мат. Сборник, **16** (1945), pp. 129–162.

[9] J. Łoś, Abelian groups that are direct summands of every abelian group which contains them as pure subgroups, *Fundamenta Math.*, **44** (1957), pp. 84–90.

[10] W. Specht, *Gruppentheorie* (Berlin—Göttingen—Heidelberg, 1956).

GENERALIZED ASSOCIATIVITY AND BISYMMETRY ON QUASIGROUPS

By

J. ACZÉL (Debrecen), V. D. BELOUSOV (Belts, USSR), and M. HOSSZÚ (Miskolc)

(*Presented by* A. RÉNYI)

§ 1. Introduction

The functional equations

$$(x\,1\,y)\,2\,z = x\,3\,(y\,4\,z), \tag{1}$$

$$(x\,\mathbf{1}\,y)\,\mathbf{2}\,(u\,\mathbf{3}\,v) = (x\,\mathbf{4}\,u)\,\mathbf{5}\,(y\,\mathbf{6}\,v) \tag{2}$$

generalize those of the associativity and bisymmetry, respectively (see [2] resp. [1]), but here one equation contains 4 and 6 unknown functions (operations), respectively.

Special cases of (1) and (2) were solved by elementary considerations in [3] and [8], respectively. Under stronger (differentiability) conditions both equations were solved in [10] and [9], respectively. F. RADÓ [13] gave a complete solution in the case where the operations (functions) figuring in these equations were defined for real variables and were strictly monotonic and continuous.

In his abstract [5] one of us has announced without proof the result containing the germ of a general solution of (1) that for a set Q forming 4 quasigroups under the operations $x\,i\,y$ ($i = 1, 2, 3, 4$) which fulfil the equation (1), there always exists an operation under which it forms a group with which all these 4 quasigroups are isotopic. (For the algebraic concepts quasigroup, isotopy etc. see e. g. [7].)

In the present paper we give the general solutions of (1) and (2) on quite arbitrary quasigroups (containing thus also a proof of the above-mentioned theorem) and some applications esp. in the theory of webs. Less complete results and proofs were already published in [12] (cf. also [14], [15]).

§ 2. Generalized associativity

We prove the following

THEOREM 1. *If a set Q forms quasigroups under all 4 operations $x\,i\,y$ ($i = 1, 2, 3, 4$) and if these operations satisfy for all elements x, y, z of this*

set the equation

$$(x\,1\,y)\,2\,z = x\,3\,(y\,4\,z), \tag{1}$$

then there exists an operation $x \circ y$ *under which* Q *forms a group with which all these* 4 *quasigroups are isotopic, in particular there exist* 5 *one-to-one mappings* $\alpha, \beta, \gamma, \delta, \varepsilon$ *of* Q *onto itself such that*

$$\begin{aligned} x\,1\,y &= (x\alpha \circ y\beta)\delta^{-1}, \quad & x\,2\,y &= x\delta \circ y\gamma, \\ x\,3\,y &= x\alpha \circ y\varepsilon, & x\,4\,y &= (x\beta \circ y\gamma)\varepsilon^{-1}. \end{aligned} \tag{3}$$

The group with which the 4 *quasigroups are isotopic is determined uniquely up to automorphisms. In the representation* (3) $x \circ y, \alpha, \beta, \gamma, \delta, \varepsilon$ *are unique up to translations by constants. — On the other hand, all isotopes of an arbitrary group built with the aid of the formulas* (3) *are quasigroups whose operations satisfy equation* (1).

PROOF OF THEOREM 1. The last assertion of the theorem becomes evident by putting the operations (3) in (1) and taking into account the associativity of $x \circ y$.

In order to prove the first assertion of the theorem we introduce the notations

$$x\varrho_i = x\,i\,a, \quad x\lambda_i = a\,i\,x, \qquad (i = 1, 2, 3, 4) \tag{4}$$

where a is an arbitrary constant (fixed element) in Q. ϱ_i^{-1}, λ_i^{-1} are the inverses of the mappings (4) which exist by the quasigroup properties of Q. If we put $x = z = a$ in (1), we have

$$y\lambda_1\varrho_2 = y\varrho_4\lambda_3. \tag{5}$$

Now by the substitutions $x = a, y = u\varrho_2^{-1}\lambda_1^{-1}, z = v\lambda_3^{-1}\lambda_4^{-1}$ and $x = u\varrho_2^{-1}\varrho_1^{-1}$, $y = a, z = v\lambda_2^{-1}\lambda_4^{-1}$ and $x = u\varrho_2^{-1}\varrho_1^{-1}, y = v\lambda_3^{-1}\lambda_4^{-1}, z = a$ in (1) we get the equations

$$u\varrho_2^{-1}\,2\,v\lambda_3^{-1}\lambda_4^{-1} = (u\varrho_2^{-1}\lambda_1^{-1}\,4\,v\lambda_3^{-1}\lambda_4^{-1})\lambda_3,$$

and

$$u\varrho_2^{-1}\,2\,v\lambda_3^{-1}\lambda_4^{-1} = u\varrho_2^{-1}\varrho_1^{-1}\,3\,v\lambda_3^{-1}$$

and

$$(u\varrho_2^{-1}\varrho_1^{1}\,1\,v\lambda_3^{-1}\varrho_4^{-1})\varrho_2 = u\varrho_2^{-1}\varrho_1^{-1}\,3\,v\lambda_3^{-1},$$

respectively. The last 3 equations show that all the 4 expressions figuring in them are equal. We denote this common value by

$$u \circ v.$$

Thus we get (3) by denoting

$$\alpha=\varrho_1\varrho_2,\quad \delta=\varrho_2,\quad \gamma=\lambda_4\lambda_3,\quad \varepsilon=\lambda_3$$

and (cf. (5)!)

$$\beta=\varrho_4\lambda_3=\lambda_1\varrho_2.$$

If we put (3) in (1), we see that the operation $x\circ y$ is associative and as an isotope of a quasigroup is again a quasigroup and associative quasigroups are groups, so Q forms a group under the operation $\circ$, q. e. d.

Finally, the assertion that the group with which the 4 quasigroups are isotopic is unique up to automorphisms is contained in the following Theorem 2 which shows also to what extent the mappings α, β, δ are determined by $x\,1\,y$ etc.

THEOREM 2. *Isotopic groups are isomorphic.* (Cf. [7], p. 57.) *In particular, if the groups* $(Q, \circ)$ *and* $(R, \cdot)$ *are isotopic:*

$$(x\circ y)\varphi=x\psi\cdot y\chi, \tag{6}$$

then they are isomorphic:

$$(x\circ y)\varkappa=x\varkappa\cdot y\varkappa, \tag{7}$$

and the mappings φ, ψ, χ *and* $\varkappa$ *are connected by the formulas*

$$x\varphi=a\cdot x\varkappa\cdot b,\quad x\psi=a\cdot x\varkappa,\quad x\chi=x\varkappa\cdot b \tag{8}$$

where a, b *are constants in* R.

PROOF OF THEOREM 2. Let e denote the unity of the group $(Q, \circ)$ and put in (6) $y=e$ and $x=e$, then one gets

$$x\psi=x\varphi\cdot b^{-1} \tag{9}$$

and

$$y\chi=a^{-1}\cdot y\varphi, \tag{10}$$

respectively, where $a=e\psi$, $b=e\chi$ and a^{-1}, b^{-1} are their inverses in $(R, \cdot)$. If we put these formulas in (6) back we get

$$(x\circ y)\varphi=x\varphi\cdot b^{-1}\cdot a^{-1}\cdot y\varphi,$$

and if we multiply by a^{-1} from the left and by b^{-1} from the right and denote

$$x\varkappa=a^{-1}\cdot x\varphi\cdot b^{-1}, \tag{11}$$

we have (7) and from (11), (9), (10) also (8) follows. Q. e. d.

REMARK. It is easy to see that Theorem 2 and its proof remains valid with minor changes if $(Q, \circ)$ is no group, only a quasigroup in which there exists an element e satisfying

$$(e\circ x)\circ e=e\circ(x\circ e)=e\circ x\circ e,\quad e\circ(x\circ y)\circ e=(e\circ x)\circ(y\circ e).$$

END OF THE PROOF OF THEOREM 1. If $(Q, \circ)$ and $(Q, \cdot)$ are groups with both of which the 4 quasigroups (or even one of them) are isotopic, then these two groups are by Theorem 2 automorphic and that was our assertion. In particular, *if also* A, B, D *are one-to-one mappings and*

$$(x\alpha \circ y\beta)\delta^{-1} = (xA \cdot yB)D^{-1}, \tag{12}$$

then

$$(x \circ y)\varkappa = x\varkappa \cdot y\varkappa$$

and

$$xA = a \cdot x\alpha\varkappa, \quad xB = x\beta\varkappa \cdot b, \quad xD = a \cdot x\delta\varkappa \cdot b.$$

This follows, by reducing (12) with

$$x\alpha = x', \quad y\beta = y'; \quad \delta^{-1}D = \varphi, \quad \alpha^{-1}A = \psi, \quad \beta^{-1}B = \chi$$

to (6), from (7) and (8).

If, specially,

$$x\delta \circ y\gamma = xD \cdot yC, \tag{13}$$

i. e. the isotopism of the two groups is "principal" ([7], p. 56), *then*

$$u \cdot v = u \circ d^{-1} \circ c^{-1} \circ v \tag{14}$$

and

$$xD = x\delta \circ d, \quad yC = c \circ y\gamma \tag{15}$$

where c, d *are arbitrary constants and* c^{-1}, d^{-1} *are their inverses* this time *with respect to the operation* $\circ$.

This can be derived from the above result (here

$$x\varkappa = c \circ x \circ d),$$

but also directly from (13): We put $x = eD^{-1}$ and $y = eC^{-1}$, respectively, (e is the unity with respect to $\cdot$) to get (15) and put (15) back in (13) to get (14).

Herefrom it is already easy to show that *from*

$$(x\alpha \circ y\beta)\delta^{-1} = (xA \cdot yB)D^{-1}, \quad x\delta \circ y\gamma = xD \cdot yC,$$
$$x\alpha \circ y\varepsilon = xA \cdot yE, \quad (x\beta \circ y\gamma)\varepsilon^{-1} = (xB \cdot yC)E^{-1}$$

it follows that

$$u \cdot v = u \circ d^{-1} \circ c^{-1} \circ v$$

and

$$xA = x\alpha \circ e'^{-1} \circ c \circ d, \quad xB = e' \circ x\beta \circ d,$$
$$xC = c \circ x\gamma, \quad xD = x\delta \circ d, \quad xE = e' \circ x\varepsilon$$

where c, d, e' are arbitrary constants. This gives a full description of the unicity situation in (3) and completes the proof of Theorem 1.

In [2] and [4] it was proved that all continuous real groups (even all continuous cancellation semigroups) are isomorphic to the addition group (or to an addition semigroup, respectively) of real numbers, so we have the following

COROLLARY 1. *If the real interval I forms quasigroups under 4 operations $x\, i\, y$ ($i = 1, 2, 3, 4$), if e. g. $x\,1\,y$, $x\,2\,y$ are continuous and if these operations satisfy for all real values x, y, z of this interval the equation* (1), *then all 4 quasigroups are isotopic to the addition group of real numbers. In particular, there exist 6 one-to-one mappings $\alpha, \beta, \gamma, \delta, \varepsilon, \varkappa$ of I onto itself ($\alpha, \beta, \gamma, \delta, \varepsilon, \varkappa$ are continuous, thus topological) unique up to multiplicative and additive constants such that*

$$x\,1\,y = (x\alpha + y\beta)\delta^{-1}, \quad x\,2\,y = (x\delta + y\gamma)\varkappa^{-1},$$
$$x\,3\,y = (x\alpha + y\varepsilon)\varkappa^{-1}, \quad x\,4\,y = (x\beta + y\gamma)\varepsilon^{-1}.$$

On the other hand, the operations thus defined always satisfy (1) *and I forms respective quasigroups under them. — If the mappings $\gamma, \delta, \varkappa, C, D, K$ are continuous and*

$$(x\alpha + y\beta)\delta^{-1} = (xA + yB)D^{-1}, \quad (x\delta + y\gamma)\varkappa^{-1} = (xD + yC)K^{-1},$$
$$(x\alpha + y\varepsilon)\varkappa^{-1} = (xA + yE)K^{-1}, \quad (x\beta + y\gamma)\varepsilon^{-1} = (xB + yC)E^{-1},$$

then (and only then)

$$xA = kx\alpha + c + d - e', \quad xB = kx\beta - c + e', \quad xC = kx\gamma + c,$$
$$xD = kx\delta + d, \quad xE = kx\varepsilon + e', \quad xK = kx\varkappa + c + d$$

where c, d, e', k are real constants ($k \neq 0$).

REMARKS. In Theorem 1 (at deriving (3) from (1)) and also in Corollary 1 we did not make full use of the quasigroup properties, we needed only that some of the mappings (4) have univalent inverses. If only this weaker condition is supposed, then e. g. in Corollary 1 the additive group of real numbers must be replaced by an additive semigroup containing 0.

It follows immediately from Theorem 1 that *the general solution of*

$$(x \,\mathrm{I}\, y)\,\mathrm{II}\,(u \,\mathrm{III}\, y) = x \,\mathrm{IV}\, u$$

(generalized transitivity) on quasigroups is

$$x \,\mathrm{I}\, y = (x\alpha \circ y\beta)\delta^{-1}, \quad x \,\mathrm{II}\, y = x\delta \circ y\gamma,$$
$$x \,\mathrm{III}\, y = [(y\beta)^{-1} \circ x\varepsilon]\gamma^{-1}, \quad x \,\mathrm{IV}\, y = x\alpha \circ y\varepsilon$$

where $\alpha, \beta, \gamma, \delta, \varepsilon$ are one-to-one mappings with the inverses $\alpha^{-1}, \beta^{-1}, \gamma^{-1}$,

$\delta^{-1}, \varepsilon^{-1}$, $\circ$ is a group operation and t^{-1} is the inverse of t in this group. — In fact, by putting

$$x\,1\,y = x\,\mathrm{I}\,y, \quad x\,2\,y = x\,\mathrm{II}\,y, \quad x\,3\,y = x\,\mathrm{IV}\,y,$$
$$u\,\mathrm{III}\,y = z \quad \text{equivalent to} \quad u = y\,4\,z$$

we get (1). — Of course, a corollary similar to Corollary 1 can be deduced also herefrom with $x\,\mathrm{III}\,y = (x\varepsilon - y\beta)\gamma^{-1}$ etc. (cf. [11]).

PROBLEM. Are theorems similar to Theorem 1 and Corollary 1 valid also if one considers cancellation groupoids (see [7], p. 9) instead of quasigroups?

§ 3. Generalized bisymmetry

We prove the following

THEOREM 3. *If a set Q forms quasigroups under all 6 operations $x\,\mathbf{i}\,y$ ($\mathbf{i} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}, \mathbf{6}$) and if these operations satisfy for all elements x, y, u, v of this set the equation*

$$(2) \qquad (x\,\mathbf{1}\,y)\,\mathbf{2}\,(u\,\mathbf{3}\,v) = (x\,\mathbf{4}\,u)\,\mathbf{5}\,(y\,\mathbf{6}\,v),$$

then there exists an operation $x \oplus y$ under which Q forms an abelian group to which all these 6 quasigroups are isotopic, in particular, 8 one-to-one mappings $\alpha, \beta, \gamma, \delta, \varepsilon, \varphi, \psi, \chi$ of Q onto itself exist such that

$$(16) \qquad \begin{aligned} &x\,\mathbf{1}\,y = (x\alpha \oplus y\beta)\delta^{-1}, \quad x\,\mathbf{2}\,y = x\delta \oplus y\varphi, \quad x\,\mathbf{3}\,y = (x\chi \oplus y\gamma)\varphi^{-1}, \\ &x\,\mathbf{4}\,y = (x\alpha \oplus y\chi)\psi^{-1}, \quad x\,\mathbf{5}\,y = x\psi \oplus y\varepsilon, \quad x\,\mathbf{6}\,y = (x\beta \oplus y\gamma)\varepsilon^{-1}. \end{aligned}$$

Here $x \oplus y$, $x\alpha$, $x\beta$, $x\gamma$, $x\delta$, $x\varepsilon$, $x\varphi$, $x\psi$, $x\lambda$ are unique up to translations by constants. — On the other hand, all isotopes of an arbitrary abelian group built with the aid of the formulas (16) *are quasigroups whose operations satisfy* (2).

PROOF OF THEOREM 3. If we put $u = a$ in (2), it goes over in an equation of the form (1) with

$$x\,1\,y = x\,\mathbf{1}\,y, \quad x\,2\,y = x\,\mathbf{2}\,(a\,\mathbf{3}\,y),$$
$$x\,3\,y = (x\,\mathbf{4}\,a)\,\mathbf{5}\,y, \quad x\,4\,y = x\,\mathbf{6}\,y.$$

So we have by (3) — writing $\oplus$ instead of $\circ$ —

$$(17) \qquad x\,\mathbf{1}\,y = (x\alpha \oplus y\beta)\delta^{-1}, \quad x\,\mathbf{2}\,y = x\delta \oplus y\varphi,$$

$$(18) \qquad x\,\mathbf{5}\,y = x\psi \oplus y\varepsilon, \quad x\,\mathbf{6}\,y = (x\beta \oplus y\gamma)\varepsilon^{-1}.$$

Now we put (17) and (18) back in (2):

(19) $$x\alpha \oplus y\beta \oplus (u\,\mathbf{3}\,v)\,\varphi = (x\,\mathbf{4}\,u)\psi \oplus y\beta \oplus v\gamma.$$

With $x = y = a$ we get herefrom

(20) $$u\,\mathbf{3}\,v = (u\chi \oplus v\gamma)\,\varphi^{-1},$$

and so (19) becomes

$$x\alpha \oplus y\beta \oplus u\chi \oplus v\gamma = (x\,\mathbf{4}\,u)\psi \oplus y\beta \oplus v\gamma$$

(it was right to write here and in (19) $(x \oplus y) \oplus z = x \oplus (y \oplus z) = x \oplus y \oplus z$, as $\oplus$ is associative). Substituting herein $y = 0\beta^{-1}$ where 0 is the unity (zero) of the group $(Q, \oplus)$, we have

(21) $$x\,\mathbf{4}\,u = (x\alpha \oplus u\chi)\psi^{-1}.$$

(17), (18), (20) and (21) give together the required (16) and if we put (16) back in (2) we have — as cancellation is allowed —

$$y\beta \oplus u\chi = u\chi \oplus y\beta,$$

i. e. the group $(Q, \oplus)$ is *abelian*. This shows at the same time that if $\oplus$ is commutative, then the operations (16) always satisfy (2).

The unicity situation can be cleared similarly as in Theorem 1. One gets that *from*

$$(x\alpha \oplus y\beta)\delta^{-1} = (xA \circ yB)D^{-1}, \qquad x\delta \oplus y\varphi = xD \circ yF,$$
$$(x\chi \oplus y\gamma)\varphi^{-1} = (xH \circ yC)F^{-1}, \quad (x\alpha \oplus y\chi)\psi^{-1} = (xA \circ yH)G^{-1},$$
$$x\psi \oplus y\varepsilon = xG \circ yE, \quad (x\beta \oplus y\gamma)\varepsilon^{-1} = (xB \circ yC)E^{-1}$$

it follows that

$$u \circ v = u \oplus v \ominus d \ominus f$$

($x \ominus y$ is the inverse operation of $x \oplus y$) *and*

$$xA = x\alpha \oplus a, \quad xB = x\beta \ominus a \oplus d \oplus d \oplus f, \quad xC = x\gamma \oplus a \ominus d \oplus e, \quad xD = x\delta \oplus d,$$
$$xE = x\varepsilon \oplus e, \quad xF = x\varphi \oplus f, \quad xG = x\psi \oplus d \ominus e \oplus f,$$
$$xH = x\chi \ominus a \oplus d \oplus d \ominus e \oplus f \oplus f$$

where a, d, e, f are arbitrary constants, and conversely. This completes the proof of Theorem 3.

Also here we have a

COROLLARY 2. *If the real interval I forms continuous quasigroups under 6 operations* $x\,\mathbf{i}\,y$ $(\mathbf{i} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}, \mathbf{6})$ *and if these operations satisfy* (2) *for all real values x, y, u, v in I, then all* 6 *quasigroups are isotopic to the*

addition group of real numbers: there exist 9 *topological mappings* α, β, γ, δ, ε, φ, ψ, χ, $\varkappa$ *of* I *onto itself, unique up to multiplicative and additive constants, such that*

$$x\,\mathbf{1}\,y=(x\alpha+y\beta)\delta^{-1},\quad x\,\mathbf{2}\,y=(x\delta+y\varphi)\varkappa^{-1},\quad x\,\mathbf{3}\,y=(x\chi+y\gamma)\varphi^{-1},$$
$$x\,\mathbf{4}\,y=(x\alpha+y\chi)\psi^{-1},\quad x\,\mathbf{5}\,y=(x\psi+y\varepsilon)\varkappa^{-1},\quad x\,\mathbf{6}\,y=(x\beta+y\gamma)\varepsilon^{-1}.$$

On the other hand, the operations thus defined always satisfy (2) *and* I *forms continuous quasigroups under them.*

Of course, a similar *problem* as in connection with Theorem 1 and Corollary 1 can be posed here, too.

The special cases considered e. g. in [3], [8], [9], [10], [13] can be treated also on the basis of our Corollaries 1 and 2. For other applications see also [12]. Here we will apply our results to the theory of webs, following a suggestion of F. RADÓ (Cluj) [13].

§ 4. Application to the theory of webs

A (plane) web consists of 3 systems of curves so that two curves belonging to the same system do not intersect and two curves of two different systems intersect in exactly one point. A web is *regular* if it is homeomorphic to that consisting of 3 systems of parallel straight lines. — As two of the three systems of curves can always be carried over by homeomorphisms in perpendicular straight lines, these can be taken as lines $x=$ constant and $y=$ constant and the third system of curves as that of niveau curves of a function $F(x, y)=x\,\mathbf{2}\,y$. We shall denote the two inverse functions of this function by $x\,\mathbf{1}\,y$ and $x\,\mathbf{3}\,y$, i. e. $x\,\mathbf{2}\,y=z$ implies $x=y\,\mathbf{1}\,z$ and $y=z\,\mathbf{3}\,x$. (If we take $y=$ constant and $z=$ constant as perpendicular straight lines, then the third system of curves are the niveau curves of $y\,\mathbf{1}\,z$ etc.) The above definition shows that a web is regular if and only if there exist 3 topological mappings $\varkappa$, φ, δ such that

$$(x\,\mathbf{2}\,y)\varkappa=x\delta+y\varphi, \tag{22}$$

i. e. the quasigroup with the operation **2** is isotopic to the addition group of real numbers. — There are several geometric criteria for a web to be regular. (See [6] for all this.)

As an example we show that *the closing of the so-called Thomsen figure is necessary and sufficient for a web to be regular* (cf. [6], pp. 20, 35—40) deriving this result from our Corollary 2. — The closing of the

Thomsen figure means (see Fig. 1) that the simultaneous validity of

(23) $$x_1 \mathbf{2}\, y_2 = x_2 \mathbf{2}\, y_1 = y, \quad x_1 \mathbf{2}\, y_3 = x_3 \mathbf{2}\, y_1 = u$$

implies that of

(24) $$x_2 \mathbf{2}\, y_3 = x_3 \mathbf{2}\, y_2.$$

As from (23)

$$x_2 = y_1 \mathbf{1}\, y, \quad y_2 = y \mathbf{3}\, x_1, \quad x_3 = y_1 \mathbf{1}\, u, \quad y_3 = u \mathbf{3}\, x_1$$

follows, we get by putting this in (24) and writing $x_1 = v$, $y_1 = x$

$$(x \mathbf{1}\, y) \mathbf{2} (u \mathbf{3}\, v) = (x \mathbf{1}\, u) \mathbf{2} (y \mathbf{3}\, v).$$

This is an equation of the form (2) and so (as in the theory of webs all curves and thus the operations **1**, **2**, **3** are supposed to be continuous), by Corollary 2, (22) is valid, we have proved the sufficiency of the condition. On the other hand, if (22) holds, then (24) follows from (23) immediately and this completes the proof of our assertion.

Also other theorems on webs can be proved on the basis of our Corollaries. In this theory the solution of our problems posed above would be significant, too.

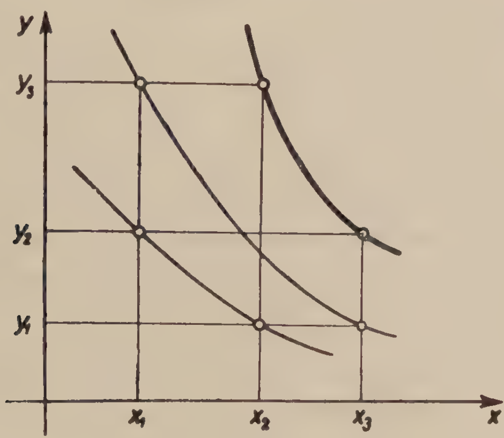


Fig. 1

The above considerations show also that the following generalization (cf. [7], p. 59 for loops instead of quasigroups) of the above theorem is true, too: *We call every quasigroup a web and regular webs those which are isotopic to abelian groups. We denote the operation in the quasigroup by* $x \mathbf{2}\, y$. *Then the implication* (23) → (24) *is necessary and sufficient for a web to be regular.*

(Received 29 September 1959)

References

[1] J. Aczél, On mean values, *Bull. Amer. Math. Soc.*, **54** (1948), pp. 392—400.

[2] J. Aczél, Sur les opérations définies pour nombres réels, *Bull. Soc. Math. de France*, 76 (1949), pp. 59—64.

[3] J. Aczél, Über einparametrige Transformationen, *Publ. Math. Debrecen*, **1** (1949—50), pp. 243—247.

[4] J. Aczél, Lösung der Vektor-Funktionalgleichung der homogenen und inhomogenen n-dimensionalen einparametrigen „Translation", der erzeugenden Funktion von Kettenreaktionen und des stationären und nicht-stationären Bewegungsintegrals, *Acta Math. Acad. Sci. Hung.*, **6** (1955), pp. 131—141.

[5] В. Д. Белоусов, Ассоциативные системы квазигрупп, Усп. Мат. Наук, **13** (1958), вып. 3 (81), p. 243.

[6] W. Blaschke und G. Bol, *Geometrie der Gewebe* (Berlin, 1938).

[7] R. H. Bruck, *A survey of binary systems* (Berlin—Göttingen—Heidelberg, 1958).

[8] M. Hosszú, A biszimmetria függvényegyenletéhez, *MTA Alk. Mat. Int. Közl.*, **1** (1952), pp. 335—342.

[9] M. Hosszú, A generalization of the functional equation of bisymmetry, *Studia Math.*, **14** (1953), pp. 100—106.

[10] M. Hosszú, Some functional equations related with the associative law, *Publ. Math. Debrecen*, **3** (1953—54), pp. 205—214.

[11] M. Hosszú, On the functional equation of transitivity, *Acta Sci. Math. Szeged*, **15** (1953—54), pp. 203—208.

[12] M. Hosszú, Belouszov egy tételéről és annak néhány alkalmazásáról, *MTA Mat. és Fiz. Oszt. Közl.*, **9** (1959), pp. 51—56.

[13] F. Radó, Ecuaţii funcţionale în legătură cu nomografia, *Studii şi Cerc. Mat. Cluj*, **9** (1958), pp. 249—319.

[14] A. Sadé, Système demosien associatif de multigroupoïdes avec un scalaire non-singulier, *Annales de la Soc. Sci. de Bruxelles Math. Astr. Phys.*, **73** (1959), pp. 231—234.

[15] A. Sadé, Entropie demosienne de multigroupoïdes et de quasigroupes, *Annales de la Soc. Sci. de Bruxelles Math. Astr. Phys.*, 73 (1959), pp. 302—309.

SOME PROBLEMS CONCERNING THE STRUCTURE OF RANDOM WALK PATHS

By

P. ERDŐS (Budapest), corresponding member of the Academy,
and S. J. TAYLOR (Birmingham)

1. Introduction. We restrict our consideration to symmetric random walk, defined in the following way. Consider the lattice formed by the points of d-dimensional Euclidean space whose coordinates are integers, and let a point $S_d(n)$ perform a move randomly on this lattice according to the rules: at time zero it is at the origin and if at any time $n-1$ $(n=1,2,\dots)$ it is at some point S of the lattice, then at time n it will be at one of the $2d$ lattice points nearest S, the probability of it being at any specified one of these being $\frac{1}{2d}$.

In the present note we examine in some detail the structure of the *path* formed by the points $S_d(0), S_d(1), \dots, S_d(n), \dots$. We will sometimes be interested in the first n points of the path, and at others in some property of the infinite path obtained as $n \to \infty$. Our results will depend to a large extent on those obtained in [2]; for convenience we shall use a notation which is consistent with that paper. In Section 2 we summarise the notations used and obtain some preliminary results which will be needed in the sequel.

The paper of DVORETZKY and ERDŐS [2] was only incidentally interested in the returns to the origin of a random walk, that is, the values of the integer n for which $S_d(n)=0$. We study these in detail in Sections 3 and 4. Since PÓLYA showed [8] as long ago as 1921 that a symmetric random walk will, with probability 1, return infinitely often to the origin if $d=1, 2$, while if $d>2$, it will wander off to infinity with probability 1, the study of returns to the origin is only interesting for $d=1$ or 2. In the case of plane random walk we obtain the asymptotic distribution of the number of returns to the origin in n steps and use these to deduce strong laws analogous to the law of the iterated logarithm. The corresponding results for the case $d=1$ were previously obtained by CHUNG and HUNT [1]. In Section 4 we examine some properties of the sequence of successive returns to the origin.

In Section 5 we consider two problems related to the behaviour of $\varrho_d(n)$, the distance from the origin of $S_d(n)$. When $d \geqq 3$, the result of PÓLYA shows that $\varrho_d(n) \to \infty$ as $n \to \infty$ and DVORETZKY and ERDŐS obtained lower

bounds for the rate at which $c_d(n)$ increases. Our first problem concerns the average rate of increase: this is of interest for any value of d and we obtain different results for the cases $d=1$, $d=2$, and $d \geqq 3$. The second problem concerns a modified form of the law of the iterated logarithm for $d \geqq 3$.

Finally, in Section 6, we consider briefly the multiplicity of points on the path. We are mainly interested in two questions: firstly, how many points of the path are entered a specified finite number of times; and secondly, how large is the maximum multiplicity occurring in a path of n steps.

We hope in a subsequent paper [5] to examine in detail the intersection properties of random walk paths.

2. Notation and preliminary results. For any fixed number of dimensions $d=1,2,\ldots$ we will be considering the space Ω_d of infinite random walks in d-space with a probability measure $\mathbf{P}(E)$ defined for measurable sets in Ω_d by extending the elementary definition of probabilities of single steps. (The measure can be defined by mapping the space of paths onto a q-adic ($q=2d$) representation of the real interval $0 \leqq x \leqq 1$, and using Lebesgue measure. Since measurability problems will not be important, we do not need to go into this.) $\mathbf{P}\{.\}$ will denote the probability that a path w in Ω_d satisfies the condition within the braces. If $E_1, E_2, \ldots, E_k, \ldots$ is a sequence of sets, then we will write

$$\mathbf{P}\{E_k \ i.\ o.\}$$

for the probability that a path w is in infinitely many of the sets E_1.

$c_1, c_2, \ldots, c_{16}$ will denote finite positive real constants. $[x]$ will denote the largest integer not greater than the real number x.

$$l_1(x) = \log_e x, \quad l_k(x) = \log \ldots \log x \qquad (k=1,2,\ldots),$$

where the logarithm is iterated k times.

ε will always denote a positive number.

If X is a vector in d-space, $|X|$ denotes the distance from X to the origin.

For paths in Ω_d, we denote by $\gamma_d(n)$ the probability that in the first $n-1$ steps, the path does not return to the origin. Clearly

$$1 = \gamma_d(1) \geqq \gamma_d(2) \geqq \cdots \geqq \gamma_d(n) \geqq \cdots > 0. \tag{2.1}$$

In [1] it is proved that, for $d \geqq 3$,

$$\gamma_d(n) \to \gamma_d > 0 \tag{2.2}$$

as $n \to \infty$, and

$$\gamma_d < \gamma_d(n) < \gamma_d + O(n^{1-d/2}); \tag{2.3}$$

for $d=2$, $\gamma_d(n)\to 0$ and the estimate found is

$$\gamma_2(n)=\frac{\pi}{\log n}+O\left(\frac{\log\log n}{(\log n)^2}\right). \tag{2.4}$$

Let us first see that (2.4) can be improved slightly to

$$\gamma_2(n)=\frac{\pi}{\log n}+O\left(\frac{1}{(\log n)^2}\right). \tag{2.5}$$

Write $u_2(r)$ for the probability that $S_2(r)=0$. Then for odd integers r, $u_2(r)=0$, while

$$u_2(2r)=\frac{1}{\pi r}+O\left(\frac{1}{r^2}\right) \quad \text{as} \quad r\to\infty. \tag{2.6}$$

Counting the last return to the origin, we have

$$\sum_{k=0}^{[n/2]}\gamma_2(n-2k)u_2(2k)=1. \tag{2.7}$$

By (2.1) this gives $\gamma_2(n)\sum_{k=0}^{[n/2]}u_2(2k)\leqq 1$, which with (2.6) gives

$$\gamma_2(n)\leqq\frac{\pi}{\log c_1 n}+O\left(\frac{1}{(\log n)^2}\right). \tag{2.8}$$

Now if $1<k_1<k_2<[n/2]$, (2.1) and (2.7) give

$$\gamma_2(n-2k_1)\sum_{k=0}^{k_1}u_2(2k)+\gamma_2(n-2k_2)\sum_{k=k_1+1}^{k_2}u_2(2k)+\sum_{k=k_2+1}^{[n/2]}u_2(2k)\geqq 1.$$

Now take $k_1=\left[\frac{n}{4}\right]$, $k_2=\left[\frac{n}{2}-\frac{n}{\log n}\right]$ and apply (2.6) and (2.8) to obtain

$$\gamma_2([n/2])\geqq\frac{\pi}{\log n}-O\left(\frac{1}{\log n}\right)^2.$$

Replacing n by $2n$ gives

$$\gamma_2(n)\geqq\frac{\pi}{\log n}-O\left(\frac{1}{\log n}\right)^2$$

which, together with (2.8) completes the proof of (2.5).

Now suppose that P is a lattice point in the plane whose distance from the origin is ϱ. Let $u_2(P,n)$ be the probability that $S_2(n)=P$. According to the position of P, it can only be reached in either an even number of steps or an odd number of steps. If it can be reached in an even number of steps and

(i) $k > \varrho^2$, then

$$u_2(P, 2k) = \frac{1}{\pi k} + \varrho^2 \cdot O\left(\frac{1}{k^2}\right); \tag{2.9}$$

(ii) while if $k < \varrho^2$, then

$$u_2(P, 2k) \leqq \left[\frac{1}{\pi k} + O\left(\frac{1}{k^2}\right)\right] e^{-\frac{\varrho^2}{2k}}. \tag{2.10}$$

The formulae (2.9) and (2.10) can be obtained by counting the number of paths (out of 4^{2k}) which end at P and using Stirling's formula, or by using the central limit theorem.

Now let $\gamma_2(P, n)$ be the probability that in the first n steps the path does not pass through P. Again assuming that P can be reached in an even number of steps we have

$$\gamma_2(P, n) + \sum_{k=1}^{[n/2]} u_2(P, 2k)\gamma_2(n-2k) = 1 \tag{2.11}$$

on considering the last return to P.

Subtracting (2.7) from (2.10) gives

$$\gamma_2(P, n) - \gamma_2(n) = \sum_{k=1}^{[n/2]} \{u_2(2k) - u_2(P, 2k)\}\gamma_2(n-2k). \tag{2.12}$$

Now suppose that

$$20 < \varrho < n^{1/3}. \tag{2.13}$$

We have, for $1 < k_1 < k_2 < [n/2]$,

$$\gamma_2(P, n) - \gamma_2(n) \leqq \gamma_2(n-k_1)\sum_{k=1}^{k_1} u_2(2k) + \gamma_2(n-k_2)\sum_{k=k_1+1}^{k_2} \{u_2(2k) - u_2(P, 2k)\} +$$
$$+ \sum_{k=k_2+1}^{[n/2]} \{u_2(2k) - u_2(P, 2k)\}.$$

Put $k_1 = \varrho^2$, $k_2 = n^{4/5}$ and use (2.1), (2.5), (2.6) and (2.10) to give

$$\gamma_2(P, n) - \gamma_2(n) \leqq \left[\frac{\pi}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right] \frac{\{\log \varrho^2 + O(1)\}}{\pi} + O(e^{-n^{1/10}})$$

or

$$\gamma_2(P, n) - \gamma_2(n) \leqq \frac{\log \varrho^2 + O(1)}{\log n}. \tag{2.14}$$

Similarly, for $1 < k_3 < [n/2]$,

$$\gamma_2(P, n) - \gamma_2(n) \geqq \gamma_2(n) \sum_{k=1}^{k_3} \{u_2(2k) - u_2(P, 2k)\}.$$

On taking $k_3 = \varrho^2/\log\log\varrho$ this gives

$$\gamma_2(P, n) - \gamma_2(n) \geqq \frac{\log \varrho^2}{\log n}\left[1 - O\left(\frac{l_3(\varrho)}{l_1(\varrho)}\right)\right]. \tag{2.15}$$

The results (2. 5), (2. 14) and (2. 15) together show that under the conditions of (2. 13)

$$\gamma_2(P, n) = \frac{2\log\varrho}{\log n}\left[1 + O\left(\frac{l_3(\varrho)}{l_1(\varrho)}\right)\right]. \tag{2.16}$$

It is trivial to show that if $\varrho \leqq 20$,

$$\gamma_2(P, n) = O\left(\frac{1}{\log n}\right). \tag{2.17}$$

Each of the results (2. 16) and (2. 17) can also be proved for points P which can only be reached in an odd number of steps from the origin: only obvious modifications to the proof are needed.

A calculation similar to the one we have carried out will show that if

$$\varrho = n^{1/2}/\psi \quad \text{and} \quad 20 < \psi < n^{1/3},$$

then

$$\gamma_2(P, n) = 1 - \frac{2\log\psi}{\log n}\left[1 + O\left(\frac{l_2(\psi)}{l_1(\psi)}\right)\right]. \tag{2.18}$$

We omit the proofs of (2. 17) and (2. 18) as these will not be needed in the sequel.

3. The number of returns to the origin. In the case $d \geqq 3$, the situation is very simple. Let γ_d be the probability that the random walk never returns to the origin. By (2. 2), $0 < \gamma_d < 1$, and if R is the total number of returns to the origin for an infinite path in d-space, the random variable R must have the geometric distribution

$$\mathbf{P}\{R = k\} = \gamma_d(1-\gamma_d)^k \qquad (k = 0, 1, 2, \ldots). \tag{3.1}$$

Let us now consider in detail the case $d = 2$ of a random walk in the plane. Let R_n be the number of returns to the origin in the first n steps. Let W_r denote the suffix of the r^{th} return to the origin. That is, there are $r-1$ returns to O among $S_2(1), S_2(2), \ldots, S_2(n-1)$ and $S_2(n) = 0$ where $n = W_r$. It is clear that

$$\gamma_2(n) = \mathbf{P}\{W_1 \geqq n\} = \mathbf{P}\{R_{n-1} = 0\}. \tag{3.2}$$

We shall see that R_n has the order of magnitude $\log n$, so let us define a new random variable T_n by

$$T_n = \frac{R_n}{\log n} \qquad (n = 3, 4, \ldots). \tag{3.3}$$

Let $x>0$ be any real number; our aim now is to try to estimate $\mathbf{P}\{T_n<x\}$: Define an integer q by

$$q=[x\log n]+1. \tag{3.4}$$

Then if $W_q<n$, we certainly have $T_n\geqq x$. That is, we have

$$\mathbf{P}\{T_n\geqq x\}\geqq\mathbf{P}\{W_q<n\}\geqq\prod_{s=1}^{q}\mathbf{P}\left\{W_s-W_{s-1}<\frac{n}{q}\right\}=\left[\mathbf{P}\left\{W_1<\frac{n}{q}\right\}\right]^q,$$

since the variables W_s-W_{s-1} $(s=1,2,\ldots,q)$ are independent and all have the same distributions as W_1. If $p=\left[\frac{n}{q}\right]$, it follows from (3.2) that

$$\mathbf{P}\{T_n\geqq x\}\geqq[1-\gamma_2(p)]^q.$$

Thus, provided $x<\frac{\log n}{(l_2(n))^{1+\varepsilon}}$, we have, on substituting the estimate (2.5) for $\gamma_2(p)$,

$$\mathbf{P}\{T_n\geqq x\}\geqq e^{-\pi x}(1+o(1))\quad\text{as}\quad n\to\infty.$$

Note further that we have

$$\mathbf{P}\{T_n\geqq x\}\geqq e^{-\pi x}(1+o((\log n)^{-1/5})) \tag{3.5}$$

uniformly in x for $x<(\log n)^{3/4}$.

We will later also need an estimate for $\mathbf{P}\{T_n\geqq x\}$ in the case $x=k\log n$ where k is a constant. The method used above is not completely accurate in this case, but it is still sufficient to give

$$\mathbf{P}\{T_n\geqq k\log n\}\geqq\frac{e^{-k\pi\log n}}{(\log n)^{2k\pi(1+\varepsilon)}} \tag{3.6}$$

for any positive number ε.

Let us now try to obtain an upper bound corresponding to the lower bound (3.5). Let E_k $(k=1,2,\ldots,q)$ be the event that precisely k of the variables W_s-W_{s-1} are greater than or equal to n, while $q-k$ of them are less than n. Then clearly, since the events E_k are mutually exclusive,

$$\mathbf{P}\{T_n<x\}\geqq\sum_{k=1}^{q}\mathbf{P}(E_k). \tag{3.7}$$

By (3.2), we have that

$$\sum_{k=1}^{q}\mathbf{P}(E_k)=\sum_{k=1}^{q}\binom{q}{k}\{\gamma_2(n)\}^k(1-\gamma_2(n))^{q-k}=1-(1-\gamma_2(n))^q.$$

Now, using the estimate (2.5), it follows that, for $x<(\log n)^{3/4}$,

$$\mathbf{P}\{T_n<x\}\geqq 1-e^{-\pi x}(1+O(\log n)^{-1/4}) \tag{3.8}$$

uniformly in x for that range.

Thus (3. 5) and (3. 8) together show that

$$\mathbf{P}\{T_n \geqq x\} = e^{-\pi x}[1 + o(\log n)^{-1/5}] \quad \text{as} \quad n \to \infty, \tag{3.9}$$

uniformly for $x < (\log n)^{3/4}$.

If x is restricted to a fixed range, say

$$c_2 < x < c_3,$$

then a better estimate can be obtained. We have, instead of (3. 8),

$$\mathbf{P}\{T_n < x\} \geqq \{1 - e^{-\pi x}\}\left\{1 + O\left(\frac{1}{\log n}\right)\right\}.$$

In this case the estimate (3. 5) can be improved to

$$\mathbf{P}\{T_n \geqq x\} \geqq e^{-\pi x}\left\{1 + O\left(\frac{\log \log n}{\log n}\right)\right\}.$$

Thus we have

$$\mathbf{P}\{T_n \geqq x\} = e^{-\pi x}\left\{1 + O\left(\frac{\log \log n}{\log n}\right)\right\} \tag{3.10}$$

uniformly for $c_2 < x < c_3$.

We need (3. 10) to allow us to obtain a satisfactory upper bound corresponding to (3. 6).

Put $s = [k(\log n)^2]$, $t = [k \log n]$. Then

$$\mathbf{P}\{T_n \geqq k \log n\} \leqq \mathbf{P}\{W_s \leqq n\} \leqq \mathbf{P}\left\{\bigcap_{r=1}^{[\log n]} \{W_{rt} - W_{(r-1)t} \leqq n\}\right\} = [\mathbf{P}\{W_t \leqq n\}]^{[\log n]},$$

since $W_{rt} - W_{(r-1)t}$ are independent and all have the same distribution as W_t.

Now $\mathbf{P}\{W_t \leqq n\} = \mathbf{P}\{R_n < t\}$. Using the estimate (3. 10), this shows that there exists a constant c_4, depending on k such that

$$\mathbf{P}\{T_n \geqq k \log n\} \leqq e^{-\pi k \log n}(\log n)^{c_4} \tag{3.11}$$

for large enough values of n.

The detailed results obtained in this section will be needed in the sequel. Let us summarise the picture in the following

THEOREM 1. *If R_n denotes the number of returns to the origin in the first n steps of a plane random walk, then*

$$\lim_{n \to \infty} \mathbf{P}\{R_n < x \log n\} = 1 - e^{-\pi x}$$

for $x < (\log n)^{3/4}$, and the limit is approached uniformly in this range.

Thus the asymptotic value of the mean of the random variable $T_n = \dfrac{R_n}{\log n}$ is $\dfrac{1}{\pi}$. This ties up with one of the main results of [2] where

it is shown that the number of lattice points entered in n steps is $\frac{\pi n}{\log n}(1+o(1))$ with probability 1. Thus the average multiplicity of points entered must be $\frac{1}{\pi}\log n$.

REMARK. CHUNG and HUNT [1] found the result corresponding to Theorem 1 for random walk on the line. They showed that if N_n denotes the number of returns to the origin in the first n steps, then $\frac{N_n}{n^{1/2}}$ has a distribution which tends to that of $|Y|$ where Y is a normally distributed variable with mean 0 and variance 1.

We now go on to consider laws of the type of the iterated logarithm for the random variable T_n. Since the methods required are complicated and not essentially new, we will not always give complete proofs. First let us consider the small values of T_n.

THEOREM 2. *If $\varphi(x)$ decreases to zero, $\varphi(x)\log x$ increases to $+\infty$ as $x\to+\infty$, and R_n is the number of returns to the origin in the first n steps of a plane random walk, then*

$$\mathbf{P}\{R_n<\varphi(n)\log n \ \textit{i. o.}\}=0 \quad \textit{or} \quad 1,$$

according as $\int^{\infty}\frac{\varphi(x)}{x\log x}dx$ *converges or diverges.*

PROOF. The integral converges or diverges with the series $\sum_1^{\infty}\varphi_k$ where $\varphi_k=\varphi(n_k)$ and $n_k=2^{2^k}$ $(k=1,2,\ldots)$.

Suppose first that $\Sigma\varphi_k$ converges. We may assume that $\varphi(x)\geqq(\log x)^{-1/10}$, since otherwise we can replace φ by

$$\psi(x)=\max\{\varphi(x),(\log x)^{-1/10}\}$$

without upsetting the convergence of the integral. Hence by Theorem 1, for k large we have

$$\mathbf{P}\{R_{n_k}<2\varphi_k\log n_k\}<c_1\varphi_k.$$

By the Borel—Cantelli lemma, there exists with probability 1 an integer k_0 such that $R_{n_k}\geqq 2\varphi_k\log n_k$ for $k\geqq k_0$. Now if $n_{k+1}>n\geqq n_k$,

$$\frac{R_n}{\log n}\geqq\frac{1}{2}\cdot\frac{R_{n_k}}{\log n_k}\geqq\varphi_k.$$

Hence

$$\frac{R_n}{\log n}\geqq\varphi(n) \quad \text{for} \quad n\geqq n_{k_0}.$$

Now suppose that $\Sigma\varphi_k$ diverges. We need to be a little more careful as the events considered above are not independent. The conditions of Theorem 1 are satisfied so that

$$\mathbf{P}\left(W_{2^r}-W_{2^{r-1}}>e^{2^r/\varphi_r}\right)=\mathbf{P}\left(W_{2^{r-1}}>e^{2^r/\varphi_r}\right)\geqq\mathbf{P}(R_{t_r}<2^{r-1}) \tag{3.12}$$

where $t_r=[e^{2^r/\varphi_r}]$. Now

$$\mathbf{P}(R_{t_r}<2^{r-1})\geqq\mathbf{P}\left\{\frac{R_{t_r}}{\log t_r}<\frac{\varphi_r}{2}\right\}>c_2\varphi_r$$

for $r\geqq r_0$, by Theorem 1, since $t_r\to\infty$. The events $W_{2^r}-W_{2^{r-1}}>e^{2^r/\varphi_r}$ $(r=1,2,\ldots)$ are independent. Hence again appealing to the Borel—Cantelli lemma, there are infinitely many integers r for which

$$W_{2^r}>e^{2^r/\varphi_r}$$

and so

$$\frac{R_{t_r}}{\log t_r}<\varphi_r\leqq\varphi(t_r),$$

for infinitely many integers t_r. This completes the proof of the theorem.

COROLLARY. *There is probability* 1 *that, for any constant* k, $R_n<\dfrac{\log n}{k\log\log n}$ *infinitely often; but for any* $\delta>0$, *there is probability* 1 *that* $R_n>\dfrac{\log n}{(\log\log n)^{1+\delta}}$ *for all large enough* n.

Now let us go on to consider the unusually large values of R_n. For this purpose we shall find it useful to look at the sequence at the points

$$m_k=[e^{e^{k/\log k}}]\qquad(k=2,3,\ldots). \tag{3.13}$$

Suppose $\psi(x)$ is a monotonic function of x which increases to $+\infty$ as $x\to+\infty$.

Write

$$\psi_k=\psi(m_k)\qquad(k=2,3,\ldots), \tag{3.14}$$

$$\mathbf{P}\left\{\frac{\pi R_{m_k}}{\log m_k}>\psi_k\right\}=p_k.$$

THEOREM 3. *If* $\psi(x)$ *is a monotonic function which increases to* $+\infty$ *as* $x\to+\infty$ *and* $\sum_{k=2}^{\infty}e^{-\psi_k}$ *converges, where* ψ_k *is given by* (3.13), (3.14), *then*

$$\mathbf{P}\left\{\frac{\pi R_n}{\log n}>\psi(n)\ i.\,o.\right\}=0.$$

PROOF. We may assume that $\psi(n) \leqq 2 \log_3 n$, since otherwise replacing $\psi(n)$ by $\psi_1(n) = \min\{\psi(n), 2\log_3 n\}$ will not effect the convergence of $\Sigma e^{-\psi_k}$ and will only strengthen the result of the theorem.

Then by (3. 13), (3. 14) and (3. 9), a simple computation shows that, for large enough k,

$$\mathbf{P}\left\{ \frac{\pi R_{m_{k+1}}}{\log m_k} > \psi_k \right\} < e^3 \cdot e^{-\psi_k}.$$

Applying the Borel—Cantelli lemma we have, since $\Sigma e^{-\psi_k}$ converges, that there exists with probability 1, an integer k_0 such that

$$\frac{\pi R_{m_{k+1}}}{\log m_k} \leqq \psi_k \quad \text{for} \quad k \geqq k_0 .$$

Then if $m_{k+1} > n \geqq m_k$ and $k \geqq k_0$,

$$\frac{\pi R_n}{\log n} \leqq \frac{\pi R_{m_{k+1}}}{\log m_k} \leqq \psi_k \leqq \psi(n);$$

and therefore there is probability 1 that

$$\frac{\pi R_n}{\log n} > \psi(n) \text{ only finitely often.}$$

To prove the converse of Theorem 3 requires a great deal more trouble due to the independence difficulties in the application of the Borel—Cantelli lemma. We state two forms of the theorem which are almost equivalent.

THEOREM 4A. *If $\psi(x)$ is a monotonic function satisfying $\psi(x) > c_5 l_3(x)$ for some $c_5 > 0$, then*

$$\mathbf{P}\left\{ \frac{\pi R_n}{\log n} > \psi(n) \ i.\,o. \right\} = 0 \quad or \quad 1,$$

according as $\sum_{k=2}^{\infty} e^{-\psi_k}$ converges or diverges, where ψ_k is given by (3. 13), (3. 14).

THEOREM 4B. *If $\psi(x)$ is a monotonic function increasing to $+\infty$ such that $\psi(x)/\log x$ decreases to zero as $x \to +\infty$, then*

$$\mathbf{P}\left\{ \frac{\pi R_n}{\log n} > \psi(n) \ i.\,o. \right\} = 0 \quad or \quad 1,$$

according as $\int^{\infty} \frac{\psi(x)}{x \log x} e^{-\psi(x)}$ converges or diverges.

COROLLARY. *If $r>4$ is a positive integer, and*

$$\psi(x)=l_3(x)+2l_4(x)+l_5(x)+\cdots+l_r'(x)+tl_{r+1}(x),$$

then

$$\mathbf{P}\left\{\frac{\pi R_n}{\log n}>\psi(n)\ i.\,o.\right\}=0 \quad or \quad 1,$$

according as $t>1$ or $t\leqq 1$.

It is clear that for functions $\psi(x)$ which satisfy the conditions of both theorems, the two Theorems 4A, 4B are equivalent. The corollary can be deduced from either. It seems likely that the condition $\psi(x)>c_5l_3(x)$ of Theorem 4A could be relaxed, but some sort of lower bound to the rate of growth of $\psi(x)$ is necessary. A proof of Theorem 4A can be obtained by making suitable modifications to the proof of ERDŐS [4]; and a proof of Theorem 4B can be obtained by modifying the proof of CHUNG and HUNT [1]. By either method the details are extremely formidable, and we do not propose to write them down as there is only one idea needed which could be described as new. This idea will be illustrated if we give a proof for the first term only of the asymptotic expansion of the critical $\psi(x)$. This is given by

THEOREM 4C. *If $\psi(x)=c\log\log\log x$, and R_n is the number of returns to the origin in the first n steps of a plane random walk, then*

$$\mathbf{P}\left\{\frac{\pi R_n}{\log n}>\psi(n)\ i.\,o.\right\}=0 \quad or \quad 1,$$

according as $c>1$ or $c\leqq 1$.

PROOF. With $c>1$, and ψ_k defined by (3. 13), (3. 14) the series $\Sigma e^{-\psi_k}$ converges. By Theorem 3 it follows immediately that

$$\mathbf{P}\left\{\frac{\pi R_n}{\log n}>\psi(n)\ i.\,o.\right\}=0.$$

For the case $c\leqq 1$, it is sufficient to prove that

$$\mathbf{P}\left\{\frac{\pi R_n}{\log n}>l_3(n)\ i.\,o.\right\}=1. \tag{3. 15}$$

Put

$$s_k=[e^{e^{k\log k}}] \qquad (k=2,3,\ldots). \tag{3. 16}$$

Let E_k be the event

$$E_k=\left\{\frac{\pi R_{s_k}}{\log s_k}>l_3(s_k)\right\}.$$

Then, by (3.9), we have

$$\mathbf{P}(E_k) = \frac{1}{k \log k}(1 + o(1)),$$

so certainly $\Sigma \mathbf{P}(E_k)$ diverges. The points s_k are sufficiently far apart to use the simplest form of argument for overcoming independence difficulties. We can show the existence of c_6 such that

$$\mathbf{P}(E_k | E_2' E_3' \ldots E_{k-1}') > c_6 \mathbf{P}(E_k). \tag{3.17}$$

At the end of s_{k-1} steps the random walk path is certainly at a distance $< s_{k-1}$ from O. If T_k is the event that there is at least one return to the origin[1] between s_{k-1} and $s_{k-1}^{\log k}$, it follows form (2.16) that

$$\mathbf{P}(T_k | E_2' E_3' \ldots E_{k-1}') > 1 - \frac{2}{\log k} \tag{3.18}$$

for large k. If T_k occurs, the path can be started from the first return to the origin after s_{k-1} and $t_k = (s_k - s_{k-1}^{\log k})$ steps taken. Hence

$$\mathbf{P}(E_k | T_k) \geqq \mathbf{P}(Q_k) \tag{3.19}$$

where Q_k is the event that in t_k steps starting from O the number of returns is not less than $\frac{1}{\pi} l_3(s_k) \log s_k$.

Now it is clear that

$$t_k > s_k - s_k^{1/2},$$

so that $\log t_k > \log s_k(1 - s_k^{-1/2})$.

It follows from (3.9) and (3.16) that

$$\mathbf{P}(Q_k) = \frac{1}{k \log k}(1 + o(1)).$$

Combining this with (3.18) and (3.19) is sufficient to prove (3.17). This shows that

$$\mathbf{P}(E_k | E_2' E_3' \ldots E_{k-1}') \text{ diverges,}$$

so that, with probability 1, the event E_k occurs infinitely often. This completes the proof of the theorem.

4. The distribution of the returns to the origin. We have seen that $\frac{R_n}{n^{1/2}}, \frac{R_n}{\log n}$ in the cases $d = 1, 2$, respectively, each have an asymptotic distribution as $n \to \infty$. As a result these ratios do not approach a limit as

[1] This is where we introduce an idea not needed in [1] or [4].

$n \to \infty$. We first show that a suitable averaging process leads to a limit, and then show that if one only counts returns at a suitably sparse subsequence, then the number of returns has an asymptotic value.

THEOREM 5. *Suppose W_s denotes the suffix of the s^{th} return to the origin. Then there are constants c_7, c_8 such that*

(i) *if the random walk is on the line, then*

$$\lim_{n\to\infty}\left\{\frac{1}{\log n}\sum_{s=1}^{n}\frac{1}{W_s^{1/2}}\right\}=c_7 \quad \textit{with probability } 1;$$

(ii) *if the random walk is in the plane, then*

$$\lim_{n\to\infty}\left\{\frac{1}{\log n}\sum_{s=1}^{n}\frac{1}{\log W_s}\right\}=c_8 \quad \textit{with probability } 1.$$

PROOF OF (ii). For a random walk in the plane define a sequence of random variables by

$$\mu(k)=\begin{cases}\dfrac{1}{\log k} & \text{if} \quad S_2(k)=0,\\ 0 & \text{otherwise}\end{cases} \qquad (k=2,3,\ldots).$$

Using (2. 6) we see that

$$\mathcal{E}\{\mu(2k)\}=\frac{1}{\pi k\log k}+O\left(\frac{1}{k^2}\right),$$

$$\mathcal{E}\{\mu(2k+1)\}=0 \qquad (k=1,2,\ldots).$$

Hence, if we put $\nu(n)=\sum_{k=2}^{n}\mu(k)$, we obtain

$$\mathcal{E}\{\nu(n)\}=\frac{1}{\pi}\log\log n+O(1). \tag{4. 1}$$

The variance $\sigma^2\{\nu(n)\}$ may be estimated since

$$\sigma^2\leqq 2\sum_{2\leqq i\leqq j\leqq n}\frac{1}{\log i}\,\frac{1}{\log j}(\mathbf{P}\{S_2(i)=0\})\cdot$$
$$\cdot(\mathbf{P}\{S_2(j)=0\,|\,S_2(i)=0\}-\mathbf{P}\{S_2(j)=0\}).$$

Using (2. 6), a simple computation shows that

$$\sigma^2\{\nu(n)\}<c_9\log\log n \tag{4. 2}$$

for a suitable positive number c_9.

Let

$$q_k = [e^{e^{k^{10}}}] \qquad (k = 1, 2, \ldots); \tag{4. 3}$$

then $\log\log q_k = k^{10}(1+o(1))$.

By Chebyshev's inequality, using (4. 1) and (4. 2), for any $\varepsilon > 0$,

$$\mathbf{P}\left\{\left|\frac{\nu(q_k)}{\log\log q_k} - \frac{1}{\pi}\right| > \varepsilon\right\} \leqq \frac{c_{10}}{k^5}$$

for a suitable constant c_{10}. Using the Borel—Cantelli lemma, we see that

$$\lim_{k\to\infty} \frac{\nu(q_k)}{\log\log q_k} = \frac{1}{\pi} \quad \text{with probability 1.}$$

Now if $q_{k+1} > n \geqq q_k$,

$$\frac{\nu(q_{k+1})}{\log\log q_k} \geqq \frac{\nu(n)}{\log\log n} \geqq \frac{\nu(q_k)}{\log\log q_{k+1}}.$$

By (4. 3) $\dfrac{\log\log q_{k+1}}{\log\log q_k} \to 1$ as $k \to \infty$. It follows immediately that, with probability 1,

$$\lim_{n\to\infty} \frac{\nu(n)}{\log\log n} = \frac{1}{\pi}. \tag{4. 4}$$

We now show that the result (4. 4) is equivalent to (ii). By Theorem 2, we know that, with probability 1, $R_n > \log n(\log\log n)^{-3/2}$ for large n. Hence

$$W_n < e^{n(\log n)^3} \quad \text{for large } n. \tag{4. 5}$$

Similarly from Theorem 3 we can deduce that

$$W_n > e^{n/\log n} \quad \text{for large } n. \tag{4. 6}$$

Now $\sum_{s=1}^{n} \frac{1}{\log W_s} = \nu(W_n)$, so that, by (4. 5) and (4. 6),

$$\frac{1}{\log n}\nu(e^{n/\log n}) < \frac{1}{\log n}\sum_{s=1}^{n}\frac{1}{\log W_s} < \frac{1}{\log n}\nu(e^{n(\log n)^3}) \tag{4. 7}$$

for large n, with probability 1.

Since, with probability 1, both sides of (4. 7) approach the limit $1/\pi$, this completes the proof that

$$\mathbf{P}\left\{\lim_{n\to\infty}\frac{1}{\log n}\sum_{1}^{n}\frac{1}{\log W_s} = \frac{1}{\pi}\right\} = 1.$$

PROOF OF (i). Precisely the same method will work in this case, using the results of CHUNG and HUNT [1] instead of Theorems 2 and 3.

THEOREM 6. (i) *For a random walk on the line, let* $R_n(2k^2)$ *be the number of integers k for which* $S_1(2k^2)=0$ $(1 \leq k \leq n)$. *Then there is a positive number* c_{11} *such that*

$$\mathbf{P}\{\lim_{n\to\infty} R_n(2k^2)/\log n = c_{11}\} = 1.$$

(ii) *For a random walk in the plane, let* $R_n(2[k \log k])$ *be the number of integers k for which* $S_2(2[k\log k])=0$ $(1 \leq k \leq n)$. *Then there is a positive number* c_{12} *such that*

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{R_n(2[k\log k])}{\log\log n} = c_{12}\right\} = 1.$$

The proof of this theorem is very similar to that of Theorem 5, so we omit it.

Let us now ask the following question about random walks in the plane. We know that the walk returns to the origin infinitely often. However, there will be some long "gaps" when the walk does not return. How long can these gaps be? To make this precise, if $g(n)$ is a monotonic function, let us ask whether or not there are only finitely many integers n for which the path $(n, n+g(n))$ does not contain *at least* one return to the origin. We have succeeded in answering this question in the following form:

THEOREM 7. *Suppose* $f(n)$ *is a monotonic function which increases to* $+\infty$ *as* $x\to\infty$, *and let* E_n *be the event that the plane random walk path does not return to the origin between n and* $n^{f(n)}$: $\mathbf{P}\{E_n \text{ i.o}\}=0$ *or* 1, *according as the series* $\sum_{k=1}^{\infty}\frac{1}{f(2^{2^k})}$ *converges or diverges.*

PROOF. Let F_n be the event $|S_2(n)| > n^{1/4}$. Then $\mathbf{P}(F_n) > 1-n^{-1/4}$, and an application of (2. 16) will give

$$\mathbf{P}(F_n \cap E_n) > (1-n^{-1/4})\frac{2\log n^{1/4}}{f(n)\log n}(1+o(1)),$$

since the behaviour of the random walk after n depends only on its position at the n^{th} step. Hence, for large n,

$$\mathbf{P}(E_n) > \frac{1}{3}\cdot\frac{1}{f(n)}. \tag{4.8}$$

Similarly, since $|S_2(n)| \leq n$, we can apply (2. 16) to give, again for large n,

$$\mathbf{P}(E_n) < \frac{3}{f(n)}. \tag{4.9}$$

Now suppose first that $\Sigma \frac{1}{f(n_k)}$ converges where $n_k = 2^{2^k}$. Put $f_1(n) = \frac{1}{2} f(n)$, and let Q_k be the event that there is no return to the origin between n_k and $n_k^{f_1(n_k-1)}$. By (4. 9), $\Sigma \mathbf{P}(Q_k)$ converges, so, with probability 1, there is an integer K such that there is a return between n_k and $n_k^{f_1(n_k-1)}$ for all $k \geqq K$. Since $n^{f(n)} \geqq n_k^{f(n_k)} = n_{k+1}^{f_1(n_k)}$, this will imply that for $n_k \leqq n < n_{k+1}$, $k \geqq K$ there is at least one return between n and $n^{f(n)}$.

Conversely suppose that $\Sigma \frac{1}{f(n_k)}$ diverges. Because of the law of zero or one, it is sufficient to show that there exists an $\eta > 0$ such that for every integer k_1 there is an integer k_2 with

$$\mathbf{P}\left\{ \bigcup_{k=k_1}^{k_2} E_{n_k} \right\} > \eta. \tag{4. 10}$$

Since $f(n)$ is monotonic, the series $\sum_{k=1}^{\infty} \frac{1}{f(n_{5k})}$ must also diverge. Thus if k_1 is given and sufficiently large, we can certainly find k_2 such that

$$\frac{1}{20} < \sum_{k_1 \leqq 5k \leqq k_2} \frac{1}{f(n_{5k})} < \frac{1}{10}. \tag{4. 11}$$

We will show that this choice of k_2 satisfies (4. 10) with $\eta = \frac{1}{40}$. Consider the events

$$D_k = E_{n_{5k}} - E_{n_{5k}} \cap \Big[\bigcup_{5k < 5r \leqq k_2} E_{n_{5r}} \Big]. \tag{4. 12}$$

The sets D_k, for k an integer satisfying $k_1 \leqq 5k \leqq k_2$, are clearly disjoint and

$$\bigcup_{k=k_1}^{k_2} E_{n_k} \supset \bigcup_{k_1 \leqq 5k \leqq k_2} D_k,$$

so that (4. 10) will immediately follow if we can prove that, for $k_1 \leqq 5k \leqq k_2$,

$$\mathbf{P}(D_k) > \frac{1}{2} \mathbf{P}(E_{n_{5k}}). \tag{4. 13}$$

There are two cases to consider in estimating $\mathbf{P}(D_k)$: (i) $r-k$ small, (ii) $r-k$ large.

(i) If r is such that $n_{5k}^{f(n_{5k})} > n_{5r}$, then the probability of no return between $n_{5k}^{f(n_{5k})}$ and $n_{5r}^{f(n_{5r})}$ is, on using (2. 16),

$$< (2+\varepsilon) \frac{\log n_{5k}}{\log n_{5r}} \frac{f(n_{5k})}{f(n_{5r})},$$

since we know that $|S_2(n_{5k}^{f(n_{5k})})| < n_{5k}^{f(n_{5k})}$. Thus we have in this case that

$$\mathbf{P}(E_{n_{5k}} \cap E_{n_{5r}}) < \frac{3}{2^{5(r-k)}} \mathbf{P}(E_{n_{5k}}). \tag{4.14}$$

(ii) On the other hand, if $n_{5k}^{f(n_{5k})} \leqq n_{5r}$, we have by (4.9) that

$$\mathbf{P}(E_{n_{5k}} \cap E_{n_{5r}}) < \frac{3}{f(n_{5r})} \mathbf{P}(E_{n_{5k}}). \tag{4.15}$$

The two cases (4.14), (4.15) together show that

$$\mathbf{P}(E_{n_{5k}} \cap E_{n_{5r}}) < 3\left(\frac{1}{f(n_{5r})} + \frac{1}{2^{5(r-k)}}\right) \mathbf{P}(E_{n_{5k}})$$

for any $r > k$. This result applied to (4.11) and (4.12) immediately gives (4.13). This completes the proof of the theorem.

We state without proof the result for a one-dimensional walk which corresponds to Theorem 7. It can be proved by similar methods.

THEOREM 7A. *Suppose $f(n)$ is a monotonic function which increases to $+\infty$ as $x \to \infty$, and E_n is the event that a random walk path on the line does not return to the origin between n and $n\{f(n)\}^2$, then*

$$\mathbf{P}(E_n \ i.o.) = 0 \quad or \quad 1,$$

according as the series $\sum \frac{1}{f(2^k)}$ converges or diverges.

We end this section by mentioning a related problem which we have been unable to solve completely. Clearly the lattice points in any given square will eventually all be entered by a plane random walk. How quickly does this happen? More precisely we have

PROBLEM. How quickly does the function $f(n)$ need to increase so that in an infinite plane random walk, with probability 1, all the lattice points within a distance n of the origin will be entered by the walk before $f(n)$ steps except for finitely many values of n?

We can show using the methods we have discussed above that $f(n) = n^{(\log n)^{1+\varepsilon}}$ is large enough, but we have failed to get a satisfactory lower estimate and have no plausible conjecture regarding a necessary and sufficient condition for the rate of increase of $f(n)$.

5. Behaviour of the distance from the origin. For a random walk in d-space we put $\varrho_d(n) = |S_d(n)|$. Then for $d = 1$, the celebrated law of the iterated logarithm gives an upper bound to $\varrho_d(n)$ for large n. The corresponding theorem in d-space is

THEOREM 8. *For random walk in d-space*

$$\mathbf{P}\left\{\varrho_d(n) > c\sqrt{\frac{2}{d}n\log\log n}\ \ i.\,o.\right\} = 0 \quad or \quad 1,$$

according as $c > 1$ *or* $c \leqq 1$.

This result must be well-known though we have not found it stated explicitly in the literature. It can be proved by modifying the proof for the case $d = 1$.

This form of the theorem deals with the unusually large values of $\varrho_d(n)$. We may ask: how large can a sphere be for $S_d(n)$ to remain outside it for $n \geqq n_0$? This is equivalent to obtaining an upper bound for the rate of escape of $S_d(n)$. The lower bound was obtained in [2]. We will need to use

LEMMA 1. *If* $d \geqq 3$, *then for a random walk in d-space*

$$\mathbf{P}\{\varrho_d(n) < n^{1/2}(\log n)^{-2}\ i.\,o.\} = 0.$$

This is a special case of the rate of escape result of [2].

LEMMA 2. *If* $d \geqq 3$ *and we start a random walk in d-space at a distance R from O, then it will enter a sphere of centre O and radius* λR $(0 < \lambda < 1)$ *with probability*

$$p = (\lambda)^{d-2}(1 + o(1))$$

as $R \to \infty$.

This is proved for Brownian motion by DVORETZKY [3]. The random walk case follows immediately from the relationship connecting them (see [7]).

Because of the result of PÓLYA, the problem of rate of escape is only meaningful for $d \geqq 3$.

THEOREM 9. *Suppose* $c < 1$, $\varrho_d(n)$ *is the distance from O at the* n^{th} *step of a random walk in d-space,* $d \geqq 3$, *and* $\tau_d(n) = \inf_{m \geqq n} \varrho_d(n)$, *then*

$$\mathbf{P}\left\{\tau_d(n) > c\sqrt{\frac{2}{d}n\log\log n}\ \ i.\,o.\right\} = 1.$$

REMARK. Since $\tau_d(n) \leqq \varrho_d(n)$, it follows from Theorem 8 that for $c > 1$ we must have

$$\mathbf{P}\left\{\tau_d(n) > c\sqrt{\frac{2}{d}n\log\log n}\ \ i.\,o.\right\} = 0.$$

The case $c = 1$ can also be decided: in fact, by taking a great deal more care one can prove that, for any $a < 1$,

$$\mathbf{P}\left\{\tau_d(n) > \left[\frac{n}{d}\left(2\log\log n + a\,l_3(n)\right)\right]^{1/2}\ i.\,o.\right\} = 1.$$

We have been unable to obtain necessary and sufficient conditions for upper bounds to the function $\tau_d(n)$ which would correspond to the results of ERDŐS [4] and FELLER [6].

PROOF OF THEOREM 9. Let c_{13}, c_{14} satisfy

(5. 1) $$0 < c < c_{13} < c_{14} < 1.$$

Consider a single axis in d-space and let $q(n)$ be the number of steps which are taken in the direction of this axis. Let

$$Q_n = \left\{ q(n) \geqq \left[\frac{n}{d}\right] \right\}.$$

Then, since steps along the direction of each of the d axes are equally likely, we have

(5. 2) $$\mathbf{P}(Q_n) \geqq \frac{1}{2}.$$

Now, by considering only the distance from the origin in the direction of this one specified axis, we have

(5. 3) $$\mathbf{P}\left\{ \varrho_d(n) > c_{14}\left(\frac{2}{d} n \log\log n\right)^{1/2} \Big| Q_n \right\} \geqq \frac{1}{(\log n)^{c_{14}}}.$$

Further, by Lemma 2,

(5. 4) $$\mathbf{P}\left\{ \tau_d(n) > c_{13}\left(\frac{2}{d} n \log\log n\right)^{1/2} \Big| \varrho_d(n) > c_{14}\left(\frac{2}{d} n \log\log n\right)^{1/2} \right\} >$$
$$> 1 - \left(\frac{c_{13}}{c_{14}}\right)^{d-2} (1 + o(1));$$

since the required probability is that of not entering a sphere of radius $c_{13}x$ if you start from a distance $c_{14}x$ from its centre. By (5.1), (5. 2), (5. 3) and (5. 4) it follows that for a suitable c_{15} we have

(5. 5) $$\mathbf{P}\left\{ \tau_d(n) > c_{13}\left(\frac{2}{d} n \log\log n\right)^{1/2} \right\} > \frac{c_{15}}{(\log n)^{c_{14}}}.$$

In order to apply the Borel—Cantelli lemma we must replace the events in (5. 5) by suitable independent ones. Let

(5. 6) $$n_k = [e^{k^{1+\delta}}] \qquad (k = 1, 2, \ldots)$$

where

(5. 7) $$1 < 1 + \delta < c_{14}^{-1}.$$

Now put

$$\mu_k = \inf_{n_{k+1} \leqq n \leqq n_{k+2}} |S_d(n) - S_d(n_k)|. \tag{5.8}$$

Putting $n_{k+1} - n_k = t_k$, we clearly have

$$\mathbf{P}\{\mu_k > \lambda\} \geqq \mathbf{P}\{\tau_d(t_k) > \lambda\} \quad \text{for any} \quad \lambda \geqq 0.$$

Hence if we put

$$E_k = \left\{ \mu_k > c_{13} \left(\frac{2}{d} t_k \log \log t_k \right)^{1/2} \right\},$$

we have by (5. 5) and (5. 6) that

$$\mathbf{P}(E_k) > \frac{c_{15}}{k^{c_{16}}} \quad \text{where} \quad 0 < c_{16} < 1.$$

But, by (5. 8), the events $E_2, E_4, \ldots, E_{2k}, \ldots$ are independent. Hence, by Borel—Cantelli,

$$\mathbf{P}\{E_k \ i.o.\} = 1. \tag{5.9}$$

Now, by Theorem 8, there exists with probability 1 an integer k_0 such that for $k \geqq k_0$

$$\varrho_d(n_k) < 2 \left(\frac{n_k}{d} \log \log n_k \right)^{1/2}.$$

Using (5. 6) this shows that

$$\varrho_d(n_k) < n_{k+1}^{1/2}. \tag{5.10}$$

Finally, by Lemma 1, there exists with probability 1 an integer k_1 such that if $k \geqq k_1$,

$$\tau_d(n_{k+2}) \geqq \frac{n_{k+2}^{1/2}}{(\log n_{k+2})^2}.$$

Again using (5. 6), this shows that

$$\tau_d(n_{k+2}) \geqq (n_{k+1} \log \log n_{k+1})^{1/2}. \tag{5.11}$$

Now suppose $k \geqq \operatorname{Max}(k_0, k_1)$ is such that E_k occurs. Then, by (5. 8) and (5. 10),

$$\inf_{n_{k+1} \leqq n \leqq n_{k+2}} |S_d(n)| > c_{13} \left(\frac{2}{d} t_k \log \log t_k \right)^{1/2} - n_{k+1}^{1/2} > c \left(\frac{2}{d} n_{k+1} \log \log n_{k+1} \right)^{1/2}$$

for large enough k, by (5. 1), (5. 5). Now using (5. 11) we see that for such values of k we have

$$\tau_d(n_{k+1}) > c \left(\frac{2}{d} n_{k+1} \log \log n_{k+1} \right)^{1/2}.$$

This completes the proof of the theorem.

We now state a theorem regarding the average rate of growth of $\varrho_d(n)$. This is relevant for $d=1,2$ as well as $d\geqq 3$.

THEOREM 10. *For a random walk in d-space, if $\varrho_d(n)$ denotes the distance from the origin of $S_d(n)$, then there are constants λ_d such that*

(i) $$\frac{1}{\log N}\sum_{n=1}^{N}\frac{n^{-1/2}}{1+\varrho_1(n)}\to\lambda_1 \quad \textit{with probability } 1,$$

(ii) $$\frac{1}{(\log N)^2}\sum_{n=1}^{N}\frac{1}{1+\{\varrho_2(n)\}^2}\to\lambda_2 \quad \textit{with probability } 1,$$

(iii) $$\frac{1}{\log N}\sum_{n=1}^{N}\frac{1}{1+\{\varrho_d(n)\}^2}\to\lambda_d \quad \textit{with probability } 1, \quad \textit{for} \quad d=3,4,\ldots.$$

PROOF. The method of proof is very similar to that used in Theorem 5. The three cases are also very similar, so we consider only the plane case (ii).

Suppose P is a lattice point in the plane. Let $\mathbf{Y}(P,n)$ be the probability that $S_2(n)=P$. Then

(a) If $|P|<n^{1/2}/\log n$,

(5.12) $$\mathbf{P}\{\mathbf{Y}(P,n)\}=0 \quad \text{or} \quad \frac{2}{\pi n}(1+o(1)),$$

according as P cannot or can be reached in n steps. Note that for fixed n half the points can be reached and the other half cannot.

(b) If $|P|>n^{1/2}\log n$,

(5.13) $$\mathbf{P}\{\mathbf{Y}(P,n)\}=O\left(\frac{1}{n^2}\right).$$

Thus $\mathscr{E}\left(\frac{1}{1+\{\varrho_d(n)\}^2}\right)=\sum_P\frac{\mathbf{P}\{\mathbf{Y}(P,n)\}}{1+|P|^2}$. Using (5.12) and (5.13) we have

$$\sum_{|P|<\frac{n^{1/2}}{\log n}}\frac{\mathbf{P}\{\mathbf{Y}(P,n)\}}{1+|P|^2}<\mathscr{E}\left(\frac{1}{1+\{\varrho_d(n)\}^2}\right)<\sum_{|P|<n^{1/2}\log n}\frac{\mathbf{P}\{\mathbf{Y}(P,n)\}}{1+P^2}+O(1)$$

which gives

$$\frac{2}{n}(\log n-\log\log n)<\mathscr{E}\left(\frac{1}{1+\{\varrho_d(n)\}^2}\right)(1+o(1))<\frac{2}{n}(\log n+\log\log n)$$

or

(5.14) $$\mathscr{E}\left(\frac{1}{1+\{\varrho_d(n)\}^2}\right)=\frac{2\log n}{n}(1+o(1)).$$

Hence we have

$$\mathcal{E}\left\{\frac{1}{(\log N)^2}\sum_{n=1}^{N}\frac{1}{1+\{\varrho_d(n)\}^2}\right\}=\frac{2}{(\log N)^2}\sum_{n=1}^{N}\frac{\log n}{n}(1+o(1))=1+o(1)$$
$$\text{as} \quad N\to\infty.$$

A similar computation will show that the variance is small. It can be shown that

$$\sigma^2\left\{\frac{1}{(\log N)^2}\sum_{n=1}^{N}\frac{1}{1+\{\varrho_d(n)\}^2}\right\}=O\left(\frac{1}{\log N}\right).$$

The argument of Theorem 5 can be applied, proving first that the limit exists as $n\to\infty$ through the sequence $r_k=[e^{k^3}]$ and then deducing the general result.

6. Multiplicity of points on a random walk. A point P of the lattice is of multiplicity $m(P, n)$ if the random walk of n steps is at P precisely $m(P, n)$ times in the first n steps. Let us first consider how many points there are which are entered once and only once. For $d=1$, there will be 0, 1, or 2 of these, while for $d\geqq 2$ there will clearly be many. Let us consider the case $d=2$ in some detail. In a plane random walk of n steps how many points have multiplicity one?

In [2] it is shown that the probability that the k^{th} step of a random walk brings it to a point not previously entered is $\gamma_2(k)$, the same is the probability of no return to the origin in the first $k-1$ steps. It is clear therefore that the probability that at the k^{th} step a plane random walk enters a new point to which it does not return before the n^{th} step is $\gamma_2(k)\gamma_2(n-k)$.

Let $M_1(n)$ be the number of points of multiplicity 1 on a plane random walk of n steps. Then clearly

$$\mathcal{E}\{M_1(n)\}=\sum_{k=0}^{n}\gamma_2(k)\gamma_2(n-k). \tag{6. 1}$$

By (2. 1), $\mathcal{E}\{M_1(n)\}\geqq(n+1)\{\gamma_2(n)\}^2$.

Using the estimate (2. 5), we have

$$\mathcal{E}\{M_1(n)\}\geqq n\left[\frac{\pi^2}{(\log n)^2}+O\left(\frac{1}{(\log n)^3}\right)\right]. \tag{6. 2}$$

Again using (2. 1) and (2. 5) we have, if $k_1=\left[\frac{n}{(\log n)^2}\right]$, that

$$\mathcal{E}\{M_1(n)\}\leqq 2\sum_{k=0}^{k_1}\gamma_2(n-k)+\gamma_2(k_1)\sum_{k=k_1}^{n-k_1}\gamma_2(k)\leqq$$
$$\leqq O\left(\frac{n}{(\log n)^3}\right)+n\left[\frac{\pi^2}{(\log n)^2}+O\left(\frac{\log\log n}{(\log n)^3}\right)\right].$$

This together with (6. 2) shows that

$$\mathcal{E}\{M_1(n)\} = n\left[\frac{\pi^2}{(\log n)^2} + O\left(\frac{\log\log n}{(\log n)^3}\right)\right]. \tag{6.3}$$

In order to estimate the variance we need

LEMMA 3. *Let $\nu(n)$ be the probability that a plane random walk path* (i) *does not return to the origin in the first n steps and* (ii) *enters a new point at the n^{th} step. Then*

$$\nu(n) \leqq \left\{\gamma_2\left(\left[\frac{n}{2}\right]\right)\right\}^2.$$

REMARK. It is clear that $\nu(n) \sim \{\gamma_2(n)\}^2$ as $n \to \infty$, but we need only an upper bound.

PROOF. Let $q = \left[\frac{n}{2}\right]$; then the probability that a random walk path has not returned to the origin in the first $q-1$ steps is $\gamma_2(q)$. Start a path from $S_2(q)$ of length $n-q+1$ steps. The probability that the last of these steps brings the path to a point not entered since $S_2(q)$ is $\gamma_2(n-q+1) \leqq \gamma_2(q)$, by (2. 1). Now if the path is to satisfy both conditions (i) and (ii) it clearly must not return to the origin in first $q-1$ steps, and the n^{th} point must certainly be a point not entered since the q^{th} step. Thus $\nu(n) \leqq \{\gamma_2(q)\}^2$, as required.

Now let p_i be the probability that in a random walk of n steps, the i^{th} step leads to a new point of multiplicity one, and let p_{ij} be the probability that at both the i^{th} and j^{th} steps points of multiplicity one are entered. By splitting the path into three parts it follows that

$$p_{ij} \leqq \gamma_2(i)\nu(j-i)\gamma_2(n-j) \tag{6.4}$$

for $0 \leqq i < j \leqq n$. Now

$$\sigma^2\{M_1(n)\} = \sum_{i,j}(p_{ij} - p_i p_j) \leqq 2 \sum_{1 \leqq i \leqq j \leqq n}(p_{ij} - p_i p_j).$$

Using Lemma 3 and (6. 4), this gives

$$\sigma^2\{M_1(n)\} \leqq 2 \sum_{1 \leqq i \leqq j \leqq n} \gamma_2(i)\gamma_2(n-j)\left\{\left[\gamma_2\left(\frac{j-i}{2}\right)\right]^2 - \gamma_2(j)\gamma_2(n-i)\right\}. \tag{6.5}$$

The double sum in (6. 5) can be estimated by splitting it into 4 parts. Let $k_1 = \left[\frac{n}{(\log n)^5}\right]$. Since all the terms are positive and less than 1,

$$\sum_{1 \leqq i \leqq j \leqq n} \leqq \sum_{1 \leqq i \leqq k_1}\ \sum_{1 \leqq j \leqq n} + \sum_{1 \leqq i \leqq n}\ \sum_{n-k_1 \leqq j \leqq n} + \sum_{1 \leqq i \leqq n}\ \sum_{i \leqq j \leqq i+k_1} + \sum_{k_1 \leqq i \leqq n-k_1}\ \sum_{i+k_1 \leqq j \leqq n-k_1}$$

The first 3 terms are $O\left(\frac{n^2}{(\log n)^5}\right)$, and the fourth, by using (2.5), is $O\left(\frac{n^2 \log\log n}{(\log n)^5}\right)$. Thus

$$\sigma^2\{M_1(n)\} = O\left(\frac{n^2 \log\log n}{(\log n)^5}\right).$$

This variance is not quite small enough for a straightforward application of Chebyshev's inequality. However, the method used in Section 5 of [2] can be applied here with only minor modifications to show that

$$\mathbf{P}\left\{\left|M_1(n) - \frac{\pi^2 n}{(\log n)^2}\right| > \frac{\varepsilon n}{(\log n)^2}\right\} = O\left(\frac{1}{(\log n)^{1+\delta}}\right) \qquad (\delta > 0),$$

and the strong law can be deduced, as in [2], by using the sequence $t_k = [e^{k^\theta}]$ for

$$\frac{1}{1+\delta} < \theta < 1.$$

For details of the method the reader is referred to [2]. Thus we have proved

THEOREM 11. *If $M_1(n)$ is the number of points of the lattice entered once and only once in the first n steps of a plane random walk, then*

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{M_1(n)(\log n)^2}{\pi^2 n} = 1\right\} = 1.$$

REMARK. For a fixed positive integer t, a modified version of the above proof will show that the number of points of multiplicity t in the first n steps of a plane random walk is given asymptotically by the same formula

$$\frac{\pi^2 n}{(\log n)^2}.$$

A much simplified version of the same argument suffices to prove

THEOREM 12. *If t is a positive integer, and $d \geqq 3$, then the number $Q_d(t, n)$ of points which are entered by a random walk in d-space precisely t times in the first n steps is such that*

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{Q_d(t, n)}{n} = \gamma_d(1-\gamma_d)^{t-1}\right\} = 1,$$

where γ_d is the probability that the path will never return to the origin.

REMARK. This means that in $d \geqq 3$ dimensions the proportion of points entered by a random walk of n steps which have a given multiplicity agrees

with the distribution (3. 1) for the number of returns to the origin. We feel sure that this result must also be true for the plane random walk, though we have not attempted to prove it.

The result of Theorem 12 shows that, for $d \geqq 3$, most of the points entered will have small multiplicity. Let us now ask what is the largest multiplicity occurring in the first n steps of a random walk.

THEOREM 13. *Let $T_d(n)$ be the upper bound of the multiplicity of points entered in the first n steps of a d-dimensional random walk ($d \geqq 3$). Then*

$$\mathbf{P}\left\{\lim_{n\to\infty}\frac{T_d(n)}{\log n}=\lambda_d\right\}=1$$

where

$$\lambda_d=-\frac{1}{\log(1-\gamma_d)} \qquad (d=3,4,\ldots).$$

PROOF. Suppose first that $\lambda>\lambda_d$; then by (3. 1),

$$\mathbf{P}\{R_d(n)>\lambda\log n\}<(1-\gamma_d)^{\lambda\log n}.$$

There are at most n points entered: hence

$$\mathbf{P}\{T_d(n)>\lambda\log n\}<n(1-\gamma_d)^{\log n}=n^{1-\lambda/\lambda_d}.$$

Using Borel—Cantelli it follows that the event $\{T_d(n)>(\lambda_d+\varepsilon)\log n\}$ happens only finitely often for the sequence $n_k=2^k$ $(k=1,2,\ldots)$ and as a result happens for only finitely many integers n with probability 1.

There are independence difficulties in obtaining the result in the opposite direction. This time we avoid these by splitting the path into a large number of small pieces.

Let

$$u=[\log n], \quad v=\left[\frac{n}{(\log n)^2}\right].$$

Consider a piece of the path containing u^2 steps. The probability that the first point of this piece is returned to in the first u steps is

$$1-\gamma_d+O\left(\frac{1}{(\log n)^{1/2}}\right)$$

by (2. 3). Hence the probability that this first point is entered $\lambda\log n$ times in the u^2 steps is at least

$$\mu(n)=\left\{1-\gamma_d+O\left(\frac{1}{(\log n)^{1/2}}\right)\right\}^{[\lambda\log n]}$$

There are at least v such pieces which are now independent. It follows that

$$\mathbf{P}\{T_d(n) > \lambda \log n\} \geqq 1 - \{1 - \mu(n)\}^v,$$

so that

$$\mathbf{P}(T_d(n) \leqq \lambda \log n\} < \{1 - \mu(n)\}^v < e^{-n^\delta}$$

for a suitable $\delta > 0$, provided $\lambda < \lambda_d$. Hence, if $\lambda = \lambda_d$, by Borel—Cantelli, there are, with probability 1, only finitely many n for which

$$T_d(n) \leqq \lambda \log n.$$

This completes the proof of the theorem.

The problem of maximum multiplicity also has a meaning in the case $d = 2$. The method used in the proof of Theorem 13, using pieces of length $[n^{1/2}]$ and the estimates (3. 6) and (3. 11) is good enough to prove that

$$\mathbf{P}\left\{\frac{1}{4\pi} \leqq \liminf_{n\to\infty} \frac{T_2(n)}{(\log n)^2} \leqq \limsup_{n\to\infty} \frac{T_2(n)}{(\log n)^2} \leqq \frac{1}{\pi}\right\} = 1.$$

We think it likely that in fact

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{T_2(n)}{(\log n)^2} = \frac{1}{\pi}\right\} = 1,$$

though we have not succeeded in proving this.

(Received 6 October 1959)

References

[1] K. L. Chung and C. A. Hunt, On the zeros of $\sum_1^n \pm 1$, *Annals of Math.*, **50** (1949), pp. 385—400.

[2] A. Dvoretzky and P. Erdős, Some problems on random walk in space, *Proceedings of second Berkeley Symposium on Stochastic Processes*, (1950), pp. 353—367.

[3] A. Dvoretzky, Brownian motion in space and subharmonic functions (under press).

[4] P. Erdős, On the law of the iterated logarithm, *Annals of Math.*, **43** (1942), pp. 419 - 436.

[5] P. Erdős and S. J. Taylor, Some intersection properties of random walk paths, *Acta Math. Acad. Sci. Hung.*, **11** (1960) (under press).

[6] W. Feller, The general form of the so-called law of the iterated logarithm, *Trans. Amer. Math. Soc.*, **54** (1943), pp. 373—402.

[7] M. Kac, Random walk and Brownian motion, *Amer. Math. Monthly*, **54** (1947), pp. 369—391.

[8] G. Pólya, Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend der Irrfahrt im Straßennetz, *Math. Annalen*, **84** (1921), pp. 149—160.

A METHOD FOR FINDING PRECISE ERROR BOUNDS OF NUMERICAL INTEGRATION FORMULAS IN HIGHER DIMENSIONS

By

L. C. HSU (Changchun, China)

(*Presented by* P. TURÁN)

In this paper a general method for determining precise bounds of errors involved in numerical evaluation of integrals over higher dimensional regions is proposed. Our method consists in demonstrating existence of a solution for a particular system of partial differential equations. As far as certain basic idea is concerned, our method may also be regarded as a generalization of that as fully utilized by S. M. NIKOLSKI in his booklet [4]. Actually NIKOLSKI confined his attention mainly to the numerical evaluation of ordinary definite integrals.

1. For the sake of simplicity in presenting the method we shall discuss here mainly the case of cubature formulas (i. e. the case of integrating functions with two variables), since, in fact, the general case of numerical integration of functions with more than two variables can be treated in a precisely similar way.

Throughout the paper we denote by $f^{(\mu,\nu)}(x,y)$ the repeated partial derivative $\partial^{\mu+\nu}f/\partial x^{\mu}\partial y^{\nu}$, i. e.

$$f^{(\mu,\nu)}(x,y)\equiv\left(\frac{\partial}{\partial x}\right)^{\mu}\left(\frac{\partial}{\partial y}\right)^{\nu}f(x,y).$$

In particular, $f^{(0,0)}(x,y)$ will represent the function $f(x,y)$ itself.

Let $f(x,y)$ be a real valued function defined on the rectangular region $Q(a\leqq x\leqq A, b\leqq y\leqq B)$, having all partial derivatives $f^{(\mu,\nu)}(x,y)$ with orders

$$\mu=0,1,\ldots,m;\quad \nu=0,1,\ldots,n\qquad (m>1, n>1).$$

It is always assumed that $f^{(\mu,\nu)}(x,y)$ $(0\leqq\mu\leqq m-1, 0\leqq\nu\leqq n-1)$ are continuous in Q; but the partial derivatives $f^{(m,\nu)}(x,y)$ $(0\leqq\nu\leqq n-1)$, $f^{(\mu,n)}(x,y)$ $(0\leqq\mu\leqq m-1)$ and $f^{(m,n)}(x,y)$ may be piecewise continuous functions. (Here we say that $\Phi(x,y)$ is a piecewise continuous function if it is piecewise continuous with respect to each variable alone.) Since the continuity of $f^{(m,\nu)}(x,y)$ is in general not assumed, we must care of not writing $f^{(\nu,m)}(x,y)$ instead of $f^{(m,\nu)}(x,y)$ etc. (cf. E. W. HOBSON [3], p. 321).

By $W^{(m,n)}(M, M_\mu, N_\nu; Q)$ we shall mean a class of functions $\{f(x, y)\}$ in which every function $f(x, y)$ has the property as described above, and satisfies the following boundedness conditions:

$$(1) \qquad |f^{(m,n)}(x, y)| \leqq M, \quad |f^{(\mu,n)}(x, y)| \leqq M_\mu, \quad |f^{(m,\nu)}(x, y)| \leqq N_\nu$$

where M, M_μ and N_ν are all positive constants $(\mu = 0, \ldots, m-1;$ $\nu = 0, \ldots, n-1)$.

Let D be a planar region contained in Q (i.e. $D \subset Q$), and let $L(f)$ be a linear functional of f of the form

$$(2) \qquad L(f) \equiv L(f(x, y)) = \sum p_{ij} f(x_i, y_j),$$

where the right-hand side of (2) is a finite summation with double integer-indices (i, j) ranging over certain intervals, p_{ij} are given weight-coefficients and (x_i, y_j) are evaluation points chosen to be fixed within the region D. The main problem to be solved here is that of determination of the supremum

$$(3) \qquad \varepsilon(L) = \sup_f \left| \iint_D f(x, y)\, dS - \sum p_{ij} f(x_i, y_j) \right|$$

where the supremum is taken over all functions f of the function class $W^{(m,n)} \equiv W^{(m,n)}(M, M_\mu, N_\nu; Q)$. Clearly the number $\varepsilon(L) \equiv \varepsilon(L; D)$ represents the precise error bound of the integration formula

$$(4) \qquad \iint_D f(x, y)\, dS \approx \sum p_{ij} f(x_i, y_j),$$

the integrand function being allowed to vary over $W^{(m,n)}$. Moreover, the number $\varepsilon(L)$ to be determined should depend only upon $W^{(m,n)}, D, p_{ij}$ and (x_i, y_j).

2. In order to obtain an explicit formula for $\varepsilon(L)$ we have to establish two lemmas of which the first one is a simple extension of the integral remainder of Taylor's expansion, viz.

LEMMA 1. *Let $f(x, y)$ be a function of $W^{(m,n)}$ and let*

$$(5) \qquad \Phi(t, y) = \sum_{\nu=0}^{n-1} \frac{(y-b)^\nu}{\nu!} f^{(m,\nu)}(t, b), \quad \Psi(x, t) = \sum_{\mu=0}^{m-1} \frac{(x-a)^\mu}{\mu!} f^{(\mu,n)}(a, t).$$

For any integer $r \geqq 1$ define

$$(6) \qquad K_r(u) = \begin{cases} \dfrac{1}{(r-1)!} u^{r-1} & \text{for} \quad u \geqq 0, \\ 0 & \text{for} \quad u < 0. \end{cases}$$

Then we have

$$f(x, y) = \sum_{\mu=0}^{m-1} \sum_{\nu=0}^{n-1} \frac{(x-a)^\mu (y-b)^\nu}{\mu!\,\nu!} f^{(\mu,\nu)}(a, b) + R_{m,n}(x, y) \tag{7}$$

where the remainder $R_{m,n}(x, y)$ *is given by*

$$\begin{aligned} R_{m,n} = &\int_a^A K_m(x-t)\,\Phi(t, y)\,dt + \int_b^B K_n(y-t)\,\Psi(x, t)\,dt + \\ &+ \int_a^A \int_b^B K_m(x-t_1) K_n(y-t_2) f^{(m,n)}(t_1, t_2)\,dt_1\,dt_2. \end{aligned} \tag{8}$$

This lemma can easily be justified by successive applications of the Taylor formula of the form

$$g(x) - \sum_{k=0}^{r-1} \frac{(x-a)^k}{k!} g^{(k)}(a) = \frac{1}{(r-1)!} \int_a^x (x-t)^{r-1} g^{(r)}(t)\,dt \tag{9}$$

which is known to be valid for any finite function $g(x)$ defined and having continuous derivatives $g'(x), \ldots, g^{(r-1)}(x)$ and a piecewise continuous derivative $g^{(r)}(x)$ on $a \leqq x < \infty$. In fact, the last term of (8) may be expressed in the form

$$\int_b^y \frac{(y-t_2)^{n-1}}{(n-1)!} dt_2 \left[\int_a^x \frac{(x-t_1)^{m-1}}{(m-1)!} f^{(m,n)}(t_1, t_2)\,dt_1 \right] =$$

$$= \int_b^y \frac{(y-t_2)^{n-1}}{(n-1)!} f^{(0,n)}(x, t_2)\,dt_2 - \int_b^B K_n(y-t_2)\,\Psi(x, t_2)\,dt_2.$$

Again, making use of (9) successively, we may find

$$\int_b^y \frac{(y-t_2)^{n-1}}{(n-1)!} f^{(0,n)}(x, t_2)\,dt_2 = f(x, y) - \sum_{\nu=0}^{n-1} \frac{(y-b)^\nu}{\nu!} f^{(0,\nu)}(x, b) =$$

$$= f(x, y) - \sum_{\nu=0}^{n-1} \sum_{\mu=0}^{m-1} \frac{(x-a)^\mu (y-b)^\nu}{\mu!\,\nu!} f^{(\mu,\nu)}(a, b) - \int_a^A K_m(x-t)\,\Phi(t, y)\,dt$$

which obviously implies (7) with (8).

Before stating Lemma 2, let us introduce the definition: "A real bounded function $\Phi(x, y)$ defined on Q is said to have no infinitude of oscillations along the direction of x or y axis, if for each fixed y or x, the function considered as that of one variable has no infinitude of oscillations."

According to this definition we may observe that the functions defined by (19) (see § 3) actually have no infinitude of oscillations along the directions of coordinate axes.

LEMMA 2. *Let* $G_i(x), H_j(y)$ $(0 \leqq i \leqq n-1, 0 \leqq j \leqq m-1)$ *and* $F(x, y)$ *be real bounded single-valued piecewise continuous functions defined on* Q *and having no infinitude of oscillations along the directions of* x *and* y *axes. Then there is a function* $f(x, y)$ *of the class* $W^{(m,n)}(M, M_\mu, N_\nu; Q)$ *satisfying the following system of partial differential equations for almost all* x $(a \leqq x \leqq A)$, *almost all* y $(b \leqq y \leqq B)$ *and almost all* (x, y) *in* Q:

$$f^{(m,n)}(x, y) = M \cdot \operatorname{sign} F(x, y), \tag{10.1}$$

$$f^{(m,i)}(x, b) = N_i \cdot \operatorname{sign} G_i(x) \qquad (i = 0, \ldots, n-1), \tag{10.2}$$

$$f^{(j,n)}(a, y) = M_j \cdot \operatorname{sign} H_j(y) \qquad (j = 0, \ldots, m-1) \tag{10.3}$$

where sign u *is the Kronecker symbol, viz.* sign $u = +1, 0, -1$ *according as* $u > 0, u = 0, u < 0$, *respectively.*

We shall give here a constructive proof. In our proof every indefinite integral is employed in the sense of Riemann (or Lebesgue). Moreover, for brevity, a function satisfying certain equations everywhere except at a certain countable set of isolated points will be simply said to satisfy the equations. First, let us integrate (10.1) with respect to x, getting

$$\varphi_1(x, y) = \int_a^x M \cdot \operatorname{sign} F(x, y) dx.$$

Hence the function $f(x, y)$ contained in the equation

$$f^{(m-1,n)}(x, y) = \varphi_1(x, y) + M_{m-1} \cdot \operatorname{sign} H_{m-1}(y) \tag{11}$$

satisfies (10.1) and (10.3) with $j = m-1$. Now, integrate (11) with respect to x again, giving

$$\varphi_2(x, y) = \int_a^x [\varphi_1(x, y) + M_{m-1} \cdot \operatorname{sign} H_{m-1}(y)] dx.$$

We see that the function f contained in the equation

$$f^{(m-2,n)}(x, y) = \varphi_2(x, y) + M_{m-2} \cdot \operatorname{sign} H_{m-2}(y) \tag{12}$$

satisfies (10.1) and (10.3) with $j = m-1, m-2$. Clearly (12) can be integrated still with respect to x unless $m-2=0$. Thus by successive integration process we must be finally led to the equation

$$f^{(0,n)}(x, y) = \varphi_m(x, y) + M_0 \cdot \operatorname{sign} H_0(y). \tag{13}$$

By induction argument we may conclude that the function f contained in (13) must satisfy (10. 1) and (10. 3) with $j=m-1, \ldots, 1, 0$ (at all the points of continuity of F and H_j).

Now it is enough to consider (13) together with (10. 2). Proceeding as before, we may integrate (13) to get

$$\psi_1(x, y)=\int_b^y [\varphi_m(x, y)+M_0 \cdot \operatorname{sign} H_0(y)]dy. \tag{14}$$

Taking (10. 2) with $i=n-1$ and integrating successively, we have

$$\theta_1(x)=\int_a^x \int_a^{x_m} \cdots \int_a^{x_2} N_{n-1} \cdot \operatorname{sign} G_{n-1}(x_1)dx_1 \ldots dx_{m-1}dx_m.$$

Then it is easily seen that the function f contained in the equation

$$f^{(0, n-1)}(x, y)=\psi_1(x, y)+\theta_1(x) \tag{15}$$

satisfies (13) and (10. 2) with $i=n-1$. Moreover, integrating (15) with respect to y again, we have

$$\psi_2(x, y)=\int_b^y [\psi_1(x, y)+\theta_1(x)]dy. \tag{16}$$

Similarly, taking (10. 2) with $i=n-2$ and integrating with respect to x repeatedly, we have

$$\theta_2(x)=\int_a^x \int_a^{x_m} \cdots \int_a^{x_2} N_{n-2} \cdot \operatorname{sign} G_{n-2}(x_1)dx_1 \ldots dx_m.$$

It is now clear that the function f contained in

$$f^{(0, n-2)}(x, y)=\psi_2(x, y)+\theta_2(x) \tag{17}$$

satisfies (15) and (10. 2) with $i=n-2$; and so it satisfies (13) and (10. 2) with $i=n-2, n-1$.

The foregoing argument can obviously be employed repeatedly, thus finally leading to the equation

$$f^{(0, 0)}(x, y)=\psi_n(x, y)+\theta_n(x). \tag{18}$$

Here the functions $\psi_k(x, y)$ $(k=2, \ldots, n)$ are of course defined by the recurrence relations (cf. (16))

$$\psi_k(x, y)=\int_b^y [\psi_{k-1}(x, y)+\theta_{k-1}(x)]dy,$$

with $\theta_k(x)$ being defined in an obvious way (cf. the definitions of $\theta_1(x), \theta_2(x)$). Since we are actually dealing with an inductive process, it may easily be inferred from (18) and all the preceding equations that the function

$$f(x, y) \equiv \psi_n(x, y) + \theta_n(x)$$

does satisfy (13) and (10. 2) with $i = 0, 1, \ldots, n-1$ (at all the points of continuity of F, G_i and H_j). This completes the constructive proof of Lemma 2.

3. We are now able to find an explicit formula for $\varepsilon(L)$. Denote

$$(19)\quad \begin{cases} F(t_1, t_2) = \iint\limits_D K_m(x-t_1)K_n(y-t_2)\,dx\,dy - L\big(K_m(x-t_1)K_n(y-t_2)\big), \\ G_\nu(t) = \iint\limits_D K_m(x-t)\,\dfrac{(y-b)^\nu}{\nu!}\,dx\,dy - L\left(K_m(x-t)\dfrac{(y-b)^\nu}{\nu!}\right), \\ H_\mu(t) = \iint\limits_D K_n(y-t)\,\dfrac{(x-a)^\mu}{\mu!}\,dx\,dy - L\left(K_n(y-t)\dfrac{(x-a)^\mu}{\mu!}\right) \end{cases}$$

where $0 \leqq \nu \leqq n-1$, $0 \leqq \mu \leqq m-1$. Apparently all the functions defined by (19) depend only upon D and the structure of the linear functional $L(\cdot)$. It is now easy to establish the following

THEOREM. *Let the numerical integration formula* (4) *be exact for polynomials of degrees at most $m-1$ and $n-1$ in x and y, respectively. Let $\varepsilon(L)$ be defined by* (3). *Then we have*

$$(20)\quad \varepsilon(L) = \sum_{\nu=0}^{n-1} N_\nu \int_a^A |G_\nu(t)|\,dt + \sum_{\mu=0}^{m-1} M_\mu \int_b^B |H_\mu(t)|\,dt + M \int_a^A \int_b^B |F(t_1, t_2)|\,dt_1\,dt_2$$

where $F(t_1, t_2)$, $G_\nu(t)$ and $H_\mu(t)$ are defined by (19).

PROOF. For any $f \in W^{(m,n)}(M, M_\mu, N_\nu; Q)$ let us denote

$$E(f) = \left| \iint\limits_D f(x, y)\,dS - \sum p_{ij} f(x_i, y_j) \right|.$$

By Lemma 1 we see that (cf. an error analysis as sketched by P. C. HAMMER and A. W. WYMORE [2], p. 66)

$$E(f) = \left| \iint\limits_D R_{m,n}(x, y)\,dx\,dy - L(R_{m,n}(x, y)) \right| \leqq$$

$$\leqq \int_a^A \left| \iint\limits_D K_m(x-t)\,\Phi(t, y)\,dx\,dy - L(K_m(x-t)\,\Phi(t, y)) \right| dt +$$

$$+ \int_b^B \left| \iint\limits_D K_n(y-t)\,\Psi(x, t)\,dx\,dy - L(K_n(y-t)\,\Psi(x, t)) \right| dt +$$

$$+\int_a^A\int_b^B |F(t_1,t_2)f^{(m,n)}(t_1,t_2)|\,dt_1dt_2 \leqq$$

$$\leqq \sum_{\nu=0}^{n-1}\int_a^A |G_\nu(t)f^{(m,\nu)}(t,b)|\,dt+\sum_{\mu=0}^{m-1}\int_b^B |H_\mu(t)f^{(\mu,n)}(a,t)|\,dt+$$

$$+\int_a^A\int_b^B |F(t_1,t_2)f^{(m,n)}(t_1,t_2)|\,dt_1dt_2 \leqq$$

$$\leqq \sum_{\nu=0}^{n-1}N_\nu\int_a^A |G_\nu(t)|\,dt+\sum_{\mu=0}^{m-1}M_\mu\int_b^B |H_\mu(t)|\,dt+M\int_a^A\int_b^B |F(t_1,t_2)|\,dt_1dt_2.$$

On the other hand, in accordance with Lemma 2, there is a function $f_0(x, y)$ of the class $W^{(m,n)}$ satisfying (10. 1), (10. 2) and (10. 3) at all the points of continuity of $F(x, y)$, $G_i(x)$ and $H_j(y)$, so that

$$E(f_0)=\left|\iint_D f_0(x,y)\,dxdy-L(f_0)\right|=$$

$$=\sum_{\nu=0}^{n-1}\int_a^A G_\nu(t)f_0^{(m,\nu)}(t,b)\,dt+\sum_{\mu=0}^{m-1}\int_b^B H_\mu(t)f_0^{(\mu,n)}(a,t)\,dt+$$

$$+\int_a^A\int_b^B F(t_1,t_2)f_0^{(m,n)}(t_1,t_2)\,dt_1dt_2=$$

$$=\sum_{\nu=0}^{n-1}\int_a^A N_\nu|G_\nu(t)|\,dt+\sum_{\mu=0}^{m-1}\int_b^B M_\mu|H_\mu(t)|\,dt+\int_a^A\int_b^B M|F(t_1,t_2)|\,dt_1dt_2.$$

This shows that the upper bound of $E(f)$ can be really attained by taking $f=f_0$. In other words we have shown that $\varepsilon(L)=\sup E(f)=E(f_0)$. Hence the formula (20) is proved.

4. It is apparent from (20) that, in order to estimate $\varepsilon(L)$ for a given numerical integration formula (4) with $f\in W^{(m,n)}$, one needs only to compute the following set of constants:

$$\varrho=\int_a^A\int_b^B |F(t_1,t_2)|\,dt_1dt_2,\quad \eta_\nu=\int_a^A |G_\nu(t)|\,dt,\quad \varepsilon_\mu=\int_b^B |H_\mu(t)|\,dt.$$

Obviously, these integrals can only be evaluated approximately, in case the given region D is not simple (noticing (19)). However, as far as only symmetrical regions (or some special regions) are concerned, there is no neces-

sity to calculate $\varepsilon(L)$ for each of the particular regions. In fact, P. C. HAMMER and his cooperator have already given some useful theorems which express how to generate numerical integration formulas for a variety of regions (especially those of symmetrical regions) and how the error functionals are transformed (see loc. cit.). In other words, the existing theory can be applied to extending the usefulness of our explicit formula (20).

As an illustration, consider, for instance, the simplest integration formula

$$\iint_Q f(x,y)\,dx\,dy \approx \frac{1}{4}[f(0,0)+f(0,1)+f(1,0)+f(1,1)] \tag{21}$$

where Q is the square region $Q(0 \leqq x \leqq 1, 0 \leqq y \leqq 1)$. It is evident that this integration formula is exact for all polynomials $f(x,y) = \Sigma\Sigma c_{ij} x^i y^j$ $(0 \leqq i, j \leqq 1)$. Thus we may take $m = n = 2$, and we have to compute the constants ϱ, η_0, η_1, and $\varepsilon_0, \varepsilon_1$. According to (19) and (6), we easily find

$$F(t_1, t_2) = \frac{1}{4}[(1-t_1)^2(1-t_2)^2 - (1-t_1)(1-t_2)],$$

$$G_0(t) = H_0(t) = \frac{1}{2}[(1-t)^2 - (1-t)],$$

$$G_1(t) = H_1(t) = \frac{1}{4}[(1-t)^2 - (1-t)].$$

Consequently we have

$$\int_0^1\int_0^1 |F(t_1,t_2)|\,dt_1\,dt_2 = \frac{5}{144}, \quad \int_0^1 |G_0(t)|\,dt = 2\int_0^1 |G_1(t)|\,dt = \frac{1}{12}.$$

Hence in accordance with (20) the precise error bound of the numerical integration formula (21) for the function class $W_\varepsilon^{(2,2)}(M, M_\mu, N_\nu; Q)$ is found to be

$$\varepsilon(L) = M\varrho + M_0\varepsilon_0 + M_1\varepsilon_1 + N_0\eta_0 + N\eta_1 \tag{22}$$

where $\varrho = \frac{5}{144}$, $\varepsilon_0 = \eta_0 = \frac{1}{12}$, $\varepsilon_1 = \eta_1 = \frac{1}{24}$.

Making use of the "*artesian product principle*" as formulated in HAMMER—WYMORE's paper (loc. cit.), we can actually construct various numerical integration formulas of higher accuracy for the integral of (21). And, correspondingly, we may compute various sets of constants $(\varrho, \eta_\nu, \varepsilon_\mu)$ for those integration formulas constructed. However, detailed evaluation of various constants or construction of some useful numerical tables can only be done more efficiently with the aid of computers.

5. All the process of analysis presented in §§ 1—3 can be analogously extended to the case of functions of three variables. As a matter of fact, if

$f(x, y, z)$ belongs to $W^{(m, n, l)}$, where $W^{(m, n, l)}$ is a function class defined in a way precisely similar to that of $W^{(m,n)}$, then we have an expansion corresponding to (7) of Lemma 1, namely,

$$f(x, y, z) = \sum_{\mu=0}^{m-1} \sum_{\nu=0}^{n-1} \sum_{\lambda=0}^{l-1} \frac{(x-a)^{\mu}(y-b)^{\nu}(z-c)^{\lambda}}{\mu!\,\nu!\,\lambda!} f^{(\mu, \nu, \lambda)}(a, b, c) + R_{m, n, l}.$$

The explicit form of the remainder $R_{m,n,l}(x, y, z)$, consisting of seven terms (3 definite integrals, 3 double integrals and 1 triple integral), can also easily be determined by successive applications of (9). Accordingly, we have also a similar result corresponding to Lemma 2 which can finally lead to a complete proof of a formula parallel to (20). Details are little more complicated than that of the two-dimensional case and may be omitted here.

Finally, it may be worth mentioning that the explicit formula (20) can possibly be used as a tool for investigating so-called "best approximate integration formulas" of the two-dimensional case. In fact, as well known, similar problems corresponding to the one-dimensional case have already been investigated quite successfully by A. Sard [5], S. M. Nikolski and Yu. Doronin [1] etc. on making use of some explicit formulas for error bounds.

DEPARTMENT OF MATHEMATICS,
NORTH-EAST PEOPLE'S UNIVERSITY (JILIN UNIVERSITY),
CHANGCHUN, CHINA

(Received 7 October 1959)

Bibliography

[1] Ю. Я. Доронин, К вопросу о формулах механических квадратур, Сборник Научных трудов Днепропетровского инженерностроительного института, (1955), pp. 210—217.

[2] P. C. Hammer and A. W. Wymore, Numerical evaluation of multiple integrals. I, *Math. Tables Aids Comput.*, 11 (1957), pp.59—67.

[3] E. W. Hobson, *Theory of functions of a real variable* (1907), pp. 318—321.

[4] С. М. Никольский, Квадратурные формулы (Москва, 1958).

[5] A. Sard, Best approximate integration formulas, best approximation formulas, *Amer. Journ. Math.*, 71 (1949), pp. 80—91.

EIN BEITRAG ZUR KOMITANTENTHEORIE

Von

S. GOŁĄB (Kraków) und M. KUCHARZEWSKI (Katowice)

(*Vorgelegt von* G. HAJÓS)

SATZ 1. *Ist der Tensor* $T^{\lambda_1\ldots\lambda_p}{}_{\mu_1\ldots\mu_q}$ *von der Valenz* (p,q) *eine algebraische Komitante des Tensors* $g_{\lambda\mu}$ *und ist* $p+q$ *eine ungerade Zahl, so ist* $T^{\cdots}_{\cdots}$ *ein Nulltensor.*

BEWEIS. Voraussetzungsgemäß haben wir

$$\bar{T}^{k_1\ldots k_p}{}_{j_1\ldots j_q} = T^{\lambda_1\ldots\lambda_p}{}_{\mu_1\ldots\mu_q} A^{k_1}_{\lambda_1}\ldots A^{k_p}_{\lambda_p}\cdot A^{\mu_1}_{j_1}\ldots A^{\mu_q}_{j_q}, \tag{1}$$

wo

$$A^{k_i}_{\lambda_i} = \frac{\partial \xi^{k_i}}{\partial \xi^{\lambda_i}}, \quad A^{\mu_l}_{j_l} = \frac{\partial \xi^{\mu_l}}{\partial \xi^{j_l}} \qquad (i=1,\ldots,p;\ l=1,\ldots,q) \tag{2}$$

gesetzt wurde. Setzen wir insbesondere in (1)

$$A^{k_i}_{\lambda_i} = -\delta^{k_i}_{\lambda_i} \qquad (\delta \text{ Kroneckersches Symbol}), \tag{3}$$

was erlaubt ist, da

$$J = \det A^{k_i}_{\lambda_i} = (-1)^n \neq 0$$

ist. Aus (3) und aus den Relationen $A^\mu_k A^k_\nu = \delta^\mu_\nu$ folgt

$$A^{\mu_i}_{j_i} = -\delta^{\mu_i}_{j_i}. \tag{4}$$

Bei (3) geht (1) infolge (4) in

$$\bar{T}^{k_1\ldots k_p}{}_{j_1\ldots j_q} = (-1)^{p+q} T^{k_1\ldots k_p}{}_{j_1\ldots j_q} = -T^{k_1\ldots k_p}{}_{j_1\ldots j_q} \tag{5}$$

über. Andererseits haben wir aber

$$\bar{T}^{k_1\ldots k_p}{}_{j_1\ldots j_q} = f^{k_1\ldots k_p}{}_{j_1\ldots j_q}(\bar{g}_{11},\ldots,\bar{g}_{nn}) \tag{6}$$

und

$$T^{k_1\ldots k_p}{}_{j_1\ldots j_q} = f^{k_1\ldots k_p}{}_{j_1\ldots j_q}(g_{11},\ldots,g_{nn}). \tag{7}$$

Weiter haben wir

$$\bar{g}_{ik} = g_{\lambda\mu} A^\lambda_i A^\mu_k,$$

was wegen (3) die Form

$$\bar{g}_{ik} = g_{\lambda\mu}(-\delta_i^\lambda)(-\delta_k^\mu) = g_{\lambda\mu}\delta_i^\lambda \delta_k^\mu = g_{ik} \tag{8}$$

annimmt. Setzt man (8) in (6) ein, so erhält man

$$\bar{T}^{k_1 \dots k_p}{}_{j_1 \dots j_q} = f^{k_1 \dots k_p}{}_{j_1 \dots j_q}(g_{11}, \dots, g_{nn}). \tag{9}$$

(5), (7) und (9) ergeben endlich

$$f^{k_1 \dots k_p}{}_{j_1 \dots j_q}(g_{11}, \dots, g_{nn}) = 0,$$

w. z. b. w.

Satz 1 kann leicht folgendermaßen verallgemeinert werden:

Satz 2. *Ist* $p+q$ *eine ungerade Zahl,* $r+s$ *eine gerade Zahl und ist der Tensor*

$$T^{\lambda_1 \dots \lambda_p}{}_{\mu_1 \dots \mu_q}$$

eine algebraische Komitante des Tensors

$$g^{\lambda_1 \dots \lambda_r}{}_{\mu_1 \dots \mu_s},$$

so muß $T^{\cdots}_{\cdots}$ *identisch verschwinden.*

(Eingegangen am 3. November 1959.)

ÜBER DIE AUS g_{ik} BESTIMMTE KOVARIANTE ABLEITUNG

Von

A. MOÓR (Szeged)

(*Vorgelegt von* G. HAJÓS)

Einleitung

Ein interessantes und — in Bezug auf die Geometrie — sehr wichtiges Gebiet der Theorie der geometrischen Objekte ist die Theorie der kovarianten Ableitung. Der Begriff der kovarianten Ableitung wurde von J. A. SCHOUTEN und in einer allgemeineren Form von S. GOŁĄB eingeführt (vgl. [2] und [5], insb. S. 74—75). In einigen Spezialfällen wurde auch die explizite Form der kovarianten Ableitung bestimmt (vgl. [1]—[3] und [4]), aber auch in den einfachsten und vielleicht in den wichtigsten Fällen, nämlich im Falle der Vektoren und Tensoren können noch mehrere Typen der kovarianten Ableitung bestimmt werden. Der Typ der kovarianten Ableitung hängt nämlich in erster Reihe von der Wahl des Hilfsobjektes ab.

Im Paragraphen 1 unserer Arbeit [3] haben wir für Hilfsobjekt die affinen Übertragungsparameter $\Lambda_i{}^j{}_k$ gewählt und somit eine Verallgemeinerung der linearen affinen kovarianten Ableitung bekommen. Im folgenden wollen wir die kovariante Ableitung der Vektoren bestimmen, wenn für Hilfsobjekt ein in i, k symmetrischer Tensor g_{ik} und seine partiellen Ableitungen $\partial_j g_{ik}$ gewählt werden.

Die kovariante Ableitung der kontra- bzw. kovarianten Vektoren ist somit durch die folgenden Forderungen definiert:

DEFINITION. *Die kovariante Ableitung D_k eines kontra- bzw. kovarianten Vektors ξ^i bzw. η_i soll den folgenden Forderungen genügen:*

A) *Die kovariante Ableitung $D_k\xi^i$ bzw. $D_k\eta_i$ des Vektors ξ^i bzw. η_i soll ein gemischter bzw. rein kovarianter Tensor zweiter Stufe sein.*

B) *Die kovariante Ableitung $D_k\xi^i$ bzw. $D_k\eta_i$ soll von $\xi^i, \partial_k\xi^i$ bzw. $\eta_i, \partial_k\eta_i$ und außerdem vom Hilfsobjekt g_{ij} und von seinen Derivierten $\partial_k g_{ij}$ abhängig sein, wo g_{ij} einen in i, j symmetrischen Tensor darstellt.*

C) *Die kovariante Ableitung $D_k\xi^i$ bzw. $D_k\eta_i$ soll eine in den Veränderlichen $\xi^i, \partial_k\xi^i, \partial_k g_{ij}$ bzw. $\eta_i, \partial_k\eta_i, \partial_k g_{ij}$ stetige Funktion sein. In den Veränderlichen g_{ij} fordern wir dagegen die Stetigkeit nur bei solchen Wertsystemen g_{ij}, für die*

$$\operatorname{Det}(g_{ik}) \neq 0$$

besteht.

Wir bemerken noch, daß die kovariante Ableitung für

$$g \overset{\text{def}}{=} \operatorname{Det}(g_{ik}) = 0$$

überhaupt nicht definiert sein muß. In den Riemannschen Räumen mit positiv-definiter Metrik ist z. B. immer $g > 0$.

Wird der in i, j symmetrische Tensor g_{ij} als der metrische Grundtensor eines Riemannschen Raumes betrachtet, so ist die durch die Forderungen A)—C) bestimmte kovariante Ableitung eine Verallgemeinerung der in der Theorie der Riemannschen Räume benützten metrischen kovarianten Ableitung:

$$\nabla_k \xi^i \overset{\text{def}}{=} \partial_k \xi^i + \Gamma_j{}^i{}_k \xi^j \tag{0. 1}$$

bzw.

$$\nabla_k \eta_i \overset{\text{def}}{=} \partial_k \eta_i - \Gamma_i{}^j{}_k \eta_j, \tag{0. 2}$$

wo

$$\Gamma_i{}^j{}_k \overset{\text{def}}{=} \frac{1}{2} g^{jr} (\partial_i g_{rk} + \partial_k g_{ri} - \partial_r g_{ik}) \tag{0. 3}$$

bedeutet und die g^{jr} — wie gewöhnlich — durch die Gleichungen

$$g^{jr} g_{ri} = \delta_i^j \tag{0. 4}$$

festgelegt sind.

In § 1 werden wir beweisen, daß aus g_{ij} und $\partial_k g_{ij}$ im wesentlichen nur die durch (0. 3) bestimmten in i, k symmetrischen Übertragungsparameter gebildet werden können. In den Paragraphen 2 und 3 bestimmen wir die Form der durch die Forderungen A) — C) festgelegten kovarianten Ableitung; weiter werden wir das Problem untersuchen, ob in den Formeln $D_k \xi^i$ bzw. $D_k \eta_i$ der kovarianten Ableitung die in den Forderungen A)—C) vorkommenden Größen immer explizit vorkommen müssen, oder die kovariante Ableitung von manchen dieser Größen unabhängig sein kann.

Es wird sich zeigen, daß im kontravarianten Falle in der Formel von $D_k \xi^i$ alle diese Größen explizit vorkommen müssen, während im kovarianten Falle $D_k \eta_i$ nicht alle der Größen η_i, g_{ij}, $\partial_k g_{ij}$ enthalten muß.

§ 1. Die aus g_{ij} und $\partial_k g_{ij}$ gebildeten Übertragungsparameter

In diesem Paragraphen werden wir zeigen, daß aus g_{ij} und $\partial_k g_{ij}$ nur die einzigen in i, k symmetrischen Übertragungsparameter

(1. 1) $\quad \overset{i}{\Gamma}{}_i{}^j{}_k \overset{\text{def}}{=} \frac{1}{2} g^{jr} (\partial_k g_{ir} + \partial_i g_{kr} - \partial_r g_{ik}),$ (1. 1a) $\quad g^{jr} g_{rk} = \delta_k^j,$

d. h. die sogenannten Christoffelschen Symbole gebildet werden können. Die

Übertragungsparameter (1. 1) spielen — wie wir das in den nächsten Paragraphen sehen werden — auch in der allgemeinen kovarianten Ableitung eine wesentliche Rolle. Das geometrische Objekt $\Gamma_i{}^j{}_k$ ist ein Übertragungsparameter, wenn bei einer zulässigen Koordinatentransformation

$$\bar{x}^i = \bar{x}^i(x^1, x^2, \ldots, x^n)$$

seine Transformationsformel die Form

$$\overline{\Gamma}_i{}^j{}_k = A_i^r \overline{A}_s^j A_k^t \Gamma_r{}^s{}_t + A_i{}^s{}_k \overline{A}_s^j \tag{1.2}$$

mit

$$A_j^p \overset{\text{def}}{=} \frac{\partial x^p}{\partial \bar{x}^j}, \qquad \overline{A}_m^k \overset{\text{def}}{=} \frac{\partial \bar{x}^k}{\partial x^m}, \qquad A_i{}^s{}_k \overset{\text{def}}{=} \frac{\partial^2 x^s}{\partial \bar{x}^i \partial \bar{x}^k}$$

hat.

Wir beweisen den folgenden

Satz 1. *Sind die in i, k symmetrischen Übertragungsparameter $\Gamma_i{}^j{}_k$ mit der Transformationsformel* (1. 2) *Funktionen von $g_{ab} = g_{ba}$ und $\partial_c g_{ab}$, sind ferner die $\Gamma_i{}^j{}_k$ in den Veränderlichen g_{ab} und $\partial_c g_{ab}$, wo* $\mathrm{Det}(g_{ab}) \neq 0$ *ist, stetig*[1], *so haben die $\Gamma_i{}^j{}_k$ die Form* (1. 1).

Beweis. Die Bedingung, daß nämlich die $\Gamma_i{}^j{}_k$ Funktionen von g_{ab} und $\partial_c g_{ab}$ seien, bedeutet, daß $\Gamma_i{}^j{}_k$ die Form

$$\Gamma_i{}^j{}_k = F_i{}^j{}_k(g_{ab}, \partial_c g_{ab}) \tag{1.3}$$

hat, wo die Funktionen $F_i{}^j{}_k$ in den $\partial_c g_{ab}$ und in den g_{ab}, wo $\mathrm{Det}(g_{ab}) \neq 0$ ist, stetig sind.

In einem anderen Koordinatensystem $\bar{x}^i$ bekommt man die Komponenten $\overline{\Gamma}_i{}^j{}_k$ dadurch, daß man in die Formel (1. 3) statt g_{ab} und $\partial_c g_{ab}$ die Komponenten $\bar{g}_{ab}$ und $\partial_{\bar{c}} \bar{g}_{ab}$ substituiert. Es ist somit

$$\overline{\Gamma}_i{}^j{}_k = F_i{}^j{}_k(\bar{g}_{ab}(\bar{x}), \partial_{\bar{c}} \bar{g}_{ab}(\bar{x})), \qquad \partial_{\bar{c}} \overset{\text{def}}{=} \frac{\partial}{\partial \bar{x}^c}. \tag{1.4}$$

Die Transformationsformel (1. 2) der Übertragungsparameter $\Gamma_i{}^j{}_k$ gibt somit nach (1. 4) und (1. 3)

$$F_i{}^j{}_k(\bar{g}_{ab}, \partial_{\bar{c}} \bar{g}_{ab}) = A_i^r \overline{A}_s^j A_k^t F_r{}^s{}_t(g_{ab}, \partial_c g_{ab}) + A_i{}^s{}_k \overline{A}_s^j. \tag{1.5}$$

Nun ist die Transformationsformel von g_{ab}

$$\bar{g}_{ab} = A_a^p A_b^q g_{pq}. \tag{1.6}$$

[1] Für die Stetigkeit in den g_{ij} vgl. die Forderung C) in der Einleitung, die auch für $\Gamma_i{}^j{}_k$ bestehen soll.

Durch partielle Ableitung nach $\bar{x}^c$ bekommt man aus (1.6)

$$(1.7)\qquad \partial_{\bar{c}}\bar{g}_{ab}=(A_a{}^p{}_c A_b^q+A_a^p A_b{}^q{}_c)g_{pq}+A_a^p A_b^q A_c^t \partial_t g_{pq}.$$

Substituiert man $\bar{g}_{ab}$ und $\partial_{\bar{c}}\bar{g}_{ab}$ von den Formeln (1.6) und (1.7) in (1.5), so erhält man ein Funktionalgleichungssystem für die unbekannten Funktionen $F_i{}^j{}_k$. Es ist

$$(1.8)\qquad \begin{aligned} &F_i{}^j{}_k(A_a^p A_b^q g_{pq},\,(A_a{}^p{}_c A_b^q+A_a^p A_b{}^q{}_c)g_{pq}+A_a^p A_b^q A_c^t\partial_t g_{pq})=\\ &=A_i^r \bar{A}_s^j A_k^t F_r{}^s{}_t(g_{ab},\,\partial_c g_{ab})+A_i{}^s{}_k \bar{A}_s^j. \end{aligned}$$

(1.8) besteht also in einem Punkte $\underset{(0)}{x^i}$ für jedes Wertsystem der $\frac{1}{2}n^2(n+3)$ Veränderlichen $A_a^p, A_a{}^p{}_c$. Die $\bar{A}_a^j$ sind auf Grund der Relationen

$$(1.9)\qquad \bar{A}_t^k A_i^t=\delta_i^k$$

durch die A_a^r bestimmt.

Für die Veränderlichen $A_b{}^p{}_c$ und A_a^r nehmen wir jetzt die Werte

$$(1.10)\qquad A_a^r=\delta_a^r,\qquad A_a{}^p{}_b=-\frac{1}{2}g^{rp}(\partial_b g_{ar}+\partial_a g_{br}-\partial_r g_{ab});$$

aus dem Funktionalgleichungssystem (1.8) wird somit

$$(1.11)\qquad F_i{}^j{}_k(g_{ab},0)=F_i{}^j{}_k(g_{ab},\partial_c g_{ab})-\frac{1}{2}g^{jr}(\partial_k g_{ir}+\partial_i g_{kr}-\partial_r g_{ik}).$$

Beachten wir jetzt die Transformationsformel (1.5) von $F_i{}^j{}_k$, so zeigt sich, daß auf der rechten Seite unserer letzten Gleichung ein Tensor steht. Das bedeutet aber, daß auch

$$T_i{}^j{}_k(g_{ab})\overset{\text{def}}{=}F_i{}^j{}_k(g_{ab},0)$$

einen Tensor bestimmt, der allein von g_{ab} abhängig und in i, k symmetrisch ist.

Auf Grund des Satzes von S. Gołąb und M. Kucharzewski (vgl. [6], Satz 1)[2] ist aber

$$T_i{}^j{}_k(g_{ab})\equiv F_i{}^j{}_k(g_{ab},0)=0,$$

und somit drückt die Relation (1.11) im Hinblick auf (1.3) eben unseren Satz 1 aus.

[2] S. die vorangehende Arbeit von Herrn S. Gołąb und M. Kucharzewski [6].

§ 2. Die kovariante Ableitung der kontravarianten Vektoren

Für die kovariante Ableitung des kontravarianten Vektors ξ^i ist der folgende Satz gültig:

SATZ 2. *Die den Forderungen* A)—C)[3] *genügende kovariante Ableitung* $D_k\xi^i$ *eines kontravarianten Vektors* ξ^i *ist von* ξ^i, g_{ab} *und von*

$$\nabla_b \xi^a \overset{\text{def}}{=} \partial_b \xi^a + \Gamma_b{}^a{}_r \xi^r \tag{2. 1}$$

abhängig, wo $\Gamma_b{}^a{}_r$ *durch* (1. 1) *und* (1. 1a) *festgelegt ist.*

BEWEIS. Nach den Forderungen A) und B) ist

$$D_k \xi^i \overset{\text{def}}{=} F^i{}_k(\xi^a, \partial_b \xi^a, g_{ab}, \partial_c g_{ab}), \tag{2. 2}$$

wo die n^2 Funktionen F^i_k zusammen einen Tensor bilden, der von ξ^a, $\partial_b\xi^a$, $g_{ab}, \partial_c g_{ab}$ abhängig ist. In einem anderen Koordinatensystem $\bar{x}^i$ bekommt man die Komponenten $\overline{D_k\xi^i}$ dadurch, daß man in F^i_k statt $\xi^a, \partial_b\xi^a, \ldots$ die transformierten Komponenten $\bar{\xi}^a, \partial_{\bar{b}}\bar{\xi}^a, \ldots$ substituiert. Es ist also

$$\overline{D_k \xi^i} \overset{\text{def}}{=} F^i{}_k(\bar{\xi}^a, \partial_{\bar{b}} \bar{\xi}^a, \bar{g}_{ab}, \partial_{\bar{c}} \bar{g}_{ab}).$$

Da die Ableitung $D_k\xi^i$ einen gemischten Tensor bildet, so wird wegen (1. 6), (1. 7) und wegen

$$\bar{\xi}^a = \bar{A}^a_r \xi^r, \tag{2. 3}$$

$$\partial_{\bar{b}} \bar{\xi}^a = \bar{A}^a_r A^t_b \partial_t \xi^r - \bar{A}^a_p \bar{A}^s_r A_b{}^p{}_s \xi^r \tag{2. 4}$$

für die F^i_k das Funktionalgleichungssystem

$$\bar{A}^i_r A^t_k F^r{}_t(\xi^a, \partial_b \xi^a, g_{ab}, \partial_c g_{ab}) = \tag{2. 5}$$

$$= F^i{}_k(\bar{A}^a_r \xi^r, \bar{A}^a_r A^t_b \partial_t \xi^r - \bar{A}^a_p \bar{A}^s_r A_b{}^p{}_s \xi^r, A^p_a A^q_b g_{pq}, (A_a{}^p{}_c A^q_b + A^p_a A_b{}^q{}_c) g_{pq} + A^p_a A^q_b A^r_c \partial_r g_{pq})$$

bestehen. Die Formel (2. 4) haben wir durch die partielle Ableitung nach $\bar{x}^b$ von (2. 3) unter Beachtung von (1. 9) bekommen.

Setzen wir nun in die Gleichung (2. 5)

$$A^j_m = \delta^j_m, \qquad A_a{}^b{}_c = -\Gamma_a{}^b{}_c,$$

wo $\Gamma_a{}^b{}_c$ durch (1. 1) bestimmt ist, so bekommt man $\bar{A}^i_r = \delta^i_r$ und

$$F^i{}_k(\xi^a, \partial_b \xi^a, g_{ab}, \partial_c g_{ab}) = S^i{}_k(\xi^a, \nabla_b \xi^a, g_{ab}) \tag{2. 6}$$

mit

$$S^i{}_k(\xi^a, \nabla_b \xi^a, g_{ab}) \overset{\text{def}}{=} F^i{}_k(\xi^a, \nabla_b \xi^a, g_{ab}, 0). \tag{2.6a}$$

[3] Vgl. die Einleitung.

Aus den Formeln (2. 2) und (2. 6) folgt aber, daß

$$D_k\xi^i = S^i_k(\xi^a, \nabla_b\xi^a, g_{ab}) \tag{2. 7}$$

besteht, und das drückt eben den Satz 2 aus.

Die Formel (2. 7) der kovarianten Ableitung zeigt, daß in der allgemeinsten Form von $D_k\xi^i$ die fundamentale kovariante Ableitung $\nabla_b\xi^a$ vorkommt. Es entsteht aber die Frage, ob die Veränderlichen $\xi^a, g_{ab}, \partial_c g_{ab}$ von der Formel von $D_k\xi^i$ nicht hinausfallen können. Im kovarianten Falle, wie wir das im nächsten Paragraphen sehen werden, müssen die entsprechenden Größen in der Formel der kovarianten Ableitung nicht alle vorkommen; im kontravarianten Falle besteht aber der

Satz 3. *Hängt die kovariante Ableitung $D_k\xi^i$ von den $\partial_b\xi^a$ ab, so muß sie auch von ξ^a, g_{ab} und $\partial_c g_{ab}$ abhängen.*

Beweis. Erstens werden wir zeigen, daß die kovariante Ableitung von $\partial_c g_{ab}$ nicht unabhängig sein kann. Wäre nämlich die kovariante Ableitung (2. 2) von $\partial_c g_{ab}$ unabhängig, so müßte aus (2. 5) durch die Substitution $A^r_k = \delta^r_k, \bar{A}^r_k = \delta^r_k$ statt (2. 6) die Relation

$$D_k\xi^i \equiv F^i_k(\xi^a, \partial_b\xi^a, g_{ab}) = F^i_k(\xi^a, \partial_b\xi^a - A_b{}^a{}_s\xi^s, g_{ab})$$

folgen. $A_b{}^a{}_s\xi^s$ bestimmt aber n^2 unabhängige Veränderlichen, da in einem Punkte $\underset{(0)}{x^i}$ diese Größen beliebig vorgegeben werden können. Das bedeutet aber, daß F^i_k neben $\partial_c g_{ab}$ auch von den $\partial_b\xi^a$ unabhängig ist; $D_k\xi^i$ wäre somit keine kovariante Ableitung.

Zweitens zeigen wir, daß die kovariante Ableitung $D_k\xi^i$ von g_{ab} nicht unabhängig sein kann. Widrigenfalls würde nämlich aus (2. 5) durch die Substitution $A^r_k = \delta^r_k$ die Relation

$$D_k\xi^i \equiv F^i_k(\xi^a, \partial_b\xi^a, \partial_c g_{ab}) = F^i_k(\xi^a, \partial_b\xi^a - A_b{}^a{}_r\xi^r, (A_a{}^p{}_c\delta^q_b + A_b{}^q{}_c\delta^p_a)g_{pq} + \partial_c g_{ab})$$

bestehen. Da F^i_k in den $\partial_c g_{ab}$ nach der Forderung C) stetig ist, wird diese Gleichung auch für $g_{ab} \to 0$ bestehen. Dies ist aber unmöglich, da in diesem Falle in der letzten Formel die n^2 Größen $x^b_a \overset{\text{def}}{=} A_r{}^b{}_a\xi^r$ nur auf der rechten Seite vorkommen, und durch die Substitution $x^a_b = \partial_b\xi^a$ die Funktion F^i_k, d. h. die kovariante Ableitung $D_k\xi^i$ von $\partial_b\xi^a$ unabhängig wäre.

Endlich zeigen wir, daß die kovariante Ableitung $D_k\xi^i$ von ξ^a nicht unabhängig sein kann. Wäre das der Fall, so würde nach (2. 5) durch die Substitution $A^a_t = \delta^a_t, \xi^t = 0$ und $A_a{}^b{}_c = -\Gamma_a{}^b{}_c$ das Gleichungssystem

$$D_k\xi^i \equiv F^i_k(\partial_b\xi^a, g_{ab}, \partial_c g_{ab}) = F^i_k(\partial_b\xi^a, g_{ab}, 0)$$

bestehen. Die kovariante Ableitung wäre somit auch von $\partial_c g_{ab}$ unabhängig, was aber — wie wir das schon gezeigt haben — nicht möglich ist.

Aus $\xi^i, \nabla_k \xi^i$ und g_{ij} können schon mehrere kovariante Ableitungen gebildet werden. Z. B.:

$$D_k \xi^i = \nabla_k \xi^i - \xi^i \xi^r g_{rk}, \quad D_k \xi^i = \nabla_k \xi^r g_{rt} \xi^t \xi^i, \quad D_k \xi^i = \nabla_k \xi^r \nabla_r \xi^i.$$

§ 3. Die kovariante Ableitung der kovarianten Vektoren

Es besteht der folgende

SATZ 4. *Die allgemeinste kovariante Ableitung $D_k \eta_i$ eines kovarianten Vektors η_i — falls die Forderungen* A)—C) *bestehen — ist von η_a; g_{ab} und von*

$$\nabla_b \eta_a \overset{\text{def}}{=} \partial_b \eta_a - \Gamma_a{}^r{}_b \eta_r \tag{3. 1}$$

abhängig. Die kovariante Ableitung $D_k \eta_i$ hat also die Form

$$D_k \eta_i = S_{ik}(\eta_a, \nabla_b \eta_a, g_{ab}), \tag{3. 2}$$

wo die n^2 Funktionen S_{ik} die Komponenten eines rein kovarianten Tensors zweiter Stufe bilden, und den Funktionalgleichungen

$$A_i^r A_k^t S_{rt}(\eta_a, \nabla_b \eta_a, g_{ab}) = S_{ik}(A_a^r \eta_r, A_a^r A_b^t \nabla_r \eta_t, A_a^r A_b^t g_{rt}) \tag{3. 2a}$$

Genüge leisten.

BEWEIS. Aus den Forderungen A)—C) folgt, daß die kovariante Ableitung $D_k \eta_i$ die folgende Form hat;

$$D_k \eta_i = F_{ik}(\eta_a, \partial_b \eta_a, g_{ab}, \partial_c g_{ab}). \tag{3. 3}$$

Die Funktionen F_{ik} bilden nach der Forderung A) einen kovarianten Tensor zweiter Stufe.

Die Transformationsformeln von η_a und $\partial_b \eta_a$ sind

$$\overline{\eta}_a = A_a^r \eta_r, \tag{3. 4}$$

$$\partial_{\bar{b}} \overline{\eta}_a = A_a{}^r{}_b \eta_r + A_a^r A_b^t \partial_t \eta_r. \tag{3. 5}$$

Beachtet man nun, daß einerseits $D_k \eta_i$ ein Tensor ist, zweitens, daß die transformierten $\overline{D_k \eta_i}$ dadurch entstehen, daß man in die Formel (3. 3) $\overline{\eta}_a, \partial_{\bar{b}} \overline{\eta}_a, \overline{g}_{ab}$ und $\partial_{\bar{c}} \overline{g}_{ab}$ einlegt, so bekommt man nach (3. 4), (3. 5), (1. 6) und (1. 7)

$$\begin{aligned} & A_i^r A_k^t F_{rt}(\eta_a, \partial_b \eta_a, g_{ab}, \partial_c g_{ab}) = \\ & = F_{ik}(A_a^r \eta_r, A_a{}^r{}_b \eta_r + A_a^r A_b^t \partial_t \eta_r, A_a^r A_b^t g_{rt}, (A_a{}^r{}_c A_b^t + A_b{}^r{}_c A_a^t) g_{rt} + A_a^r A_b^t A_c^s \partial_s g_{rt}). \end{aligned} \tag{3. 6}$$

(3. 6) muß in den n^2 Größen A_a^r und in den $\frac{1}{2} n^2(n+1)$ Größen $A_a{}^r{}_b$ eine Identität sein. Für $A_i^r = \delta_i^r$, $A_a{}^b{}_c = -\Gamma_c{}^b{}_a$ wird

$$D_k \eta_i \equiv F_{ik}(\eta_a, \partial_b \eta_a, g_{ab}, \partial_c g_{ab}) = F_{ik}(\eta_a, \nabla_b \eta_a, g_{ab}, 0)$$

bestehen. Mit der Bezeichnung

$$S_{ik}(\eta_a, \nabla_b \eta_a, g_{ab}) \overset{\text{def}}{=} F_{ik}(\eta_a, \nabla_b \eta_a, g_{ab}, 0)$$

bekommt man eben die Formel (3. 1), w. z. b. w.

Durch die Formeln (3. 1) und (3. 2) ist also die allgemeinste kovariante Ableitung $D_k \eta_i$ eines kovarianten Vektors η_i bestimmt, die den Forderungen A)—C) genügt, falls S_{ik} den Funktionalgleichungen (3. 2a) Genüge leistet. In der Formel (3. 2) von $D_k \eta_i$ müssen aber nicht alle Größen $\eta_a, \partial_b \eta_a, g_{ab}, \partial_c g_{ab}$ vorkommen, wie im kontravarianten Falle. Das Analogon des Satzes 3 besteht also nicht. Das zeigt sich durch den folgenden

Hilfssatz. *Aus $\partial_k \eta_i$ kann der einzige rein kovariante Tensor zweiter Stufe*

(3. 7) $$t_{ik} = c(x)\, \partial_{[k} \eta_{i]}, \qquad \partial_{[k} \eta_{i]} \overset{\text{def}}{=} \frac{1}{2} (\partial_k \eta_i - \partial_i \eta_k)$$

gebildet werden, wo der Skalar $c(x)$ beliebig gewählt werden kann.

Bemerkung. Offenbar kann der durch die Formel (3. 7) bestimmte Tensor als eine kovariante Ableitung von η_i interpretiert werden; in der Formel dieser kovarianten Ableitung kommen aber die Größen $\eta_a, g_{ab}, \partial_c g_{ab}$ nicht vor. In unserem Aufsatz [3] im Satz 8 ist die Formel (3. 7) enthalten, nur soll dort $c_1 = -c_2 = c$ gesetzt werden.

Beweis des Hilfssatzes. Bilden wir aus $\partial_b \eta_a$ einen Tensor zweiter Stufe F_{ik}, so bekommt man aus (3. 6) das analoge Funktionalgleichungssystem

$$A_i^r A_k^t F_{rt}(\partial_b \eta_a) = F_{ik}(A_a{}^r{}_b \eta_r + A_a^r A_b^t \partial_t \eta_r).$$

Durch die Substitution $A_b^t = \delta_b^t$, $A_a{}^r{}_b \eta_r = x_{ab} \equiv x_{ba}$ ergibt sich

$$F_{ik}(\partial_b \eta_a) = F_{ik}(x_{ab} + \partial_b \eta_a).$$

Nun setzen wir

(3. 8) $$x_{pq} \equiv x_{qp} = -\partial_{(p} \eta_{q)}, \qquad \partial_{(p} \eta_{q)} \overset{\text{def}}{=} \frac{1}{2} (\partial_p \eta_q + \partial_q \eta_p),$$

somit wird

$$F_{ik}(\partial_b \eta_a) = t_{ik}(\partial_{[b} \eta_{a]}),$$

wo

$$t_{ik}(\partial_{[b} \eta_{a]}) \overset{\text{def}}{=} F_{ik}\left(\frac{1}{2} (\partial_b \eta_a - \partial_a \eta_b)\right)$$

einen Tensor bedeutet, der aus $\partial_{[b}\eta_{a]}$ gebildet ist. Selbstverständlich steht in der Formel F_{ik} an der Stelle von $\partial_a \eta_a$ (nicht summieren auf a) Null.

Nach der Transformationsformel (3. 5) folgt aber unmittelbar, daß $\partial_{[b}\eta_{a]}$ ein schiefsymmetrischer rein kovarianter Tensor zweiter Stufe ist. Nach dem Satz 2 unserer Arbeit [4] hat aber dann t_{ik} die Form (3. 7), w. z. b. w.

Wir wollen nun das Problem untersuchen, was für eine Form die kovariante Ableitung $D_k\eta_i$ haben muß, wenn $D_k\eta_i$ von $\eta_a, \partial_b\eta_a$ und $\partial_c g_{ab}$ abhängig ist. Es wird sich zeigen, daß $D_k\eta_i$ von $\partial_c g_{ab}$ unabhängig sein muß. Das zeigt sich im folgenden

SATZ 5. *Ist die kovariante Ableitung $D_k\eta_i$ von g_{ab} unabhängig, so ist sie auch von $\partial_c g_{ab}$ unabhängig und $D_k\eta_i$ ist eine Funktion von η_a und $\partial_{[b}\eta_{a]}$.*

BEWEIS. Das Funktionalgleichungssystem (3. 6) wird jetzt durch die Substitution $A_a^r = \delta_a^r$, $A_a{}^r{}_b\,\eta_r = x_{ab} \equiv x_{ba}$ die Form

$$F_{ik}(\eta_a, \partial_b\eta_a, \partial_c g_{ab}) = F_{ik}(\eta_a, x_{ab} + \partial_b\eta_a, A_a{}^r{}_c g_{rb} + A_b{}^r{}_c g_{ra} + \partial_c g_{ab})$$

haben. Wenn $g_{ab} \to 0$, so wird

$$F_{ik}(\eta_a, \partial_b\eta_a, \partial_c g_{ab}) = F_{ik}(\eta_a, x_{ab} + \partial_b\eta_a, \partial_c g_{ab})$$

bestehen. $g_{ab} \to 0$ ist jetzt möglich, da F_{ik} nach der Forderung C) in den $\partial_c g_{ab}$ stetig ist. Durch Substitution (3. 8) wird F_{ik} die Form

$$D_k\eta_i \equiv F_{ik}(\eta_a, \partial_b\eta_a, \partial_c g_{ab}) = S_{ik}(\eta_a, \partial_{[b}\eta_{a]}, \partial_c g_{ab}) \tag{3. 9}$$

haben, wo S_{ik} selbstverständlich einen rein kovarianten Tensor bedeutet. Da $\partial_{[b}\eta_{a]}$ auch ein Tensor ist, wird für S_{ik} die folgende Transformationsformel bestehen:

$$A_i^r A_k^t S_{rt}(\eta_a, \partial_{[b}\eta_{a]}, \partial_c g_{ab}) =$$

$$= S_{ik}(A_a^r\eta_r, A_a^r A_b^t \partial_{[t}\eta_{r]}, (A_b^t A_a{}^r{}_c + A_a^t A_b{}^r{}_c) g_{rt} + A_a^r A_b^t A_c^s \partial_s g_{rt}).$$

Nehmen wir jetzt für A_a^r und $A_a{}^r{}_b$ die Werte

$$A_a^r = \delta_a^r, \qquad A_a{}^r{}_b = -\Gamma_a{}^r{}_b,$$

so bekommt man

$$S_{ik}(\eta_a, \partial_{[b}\eta_{a]}, \partial_c g_{ab}) \equiv S_{ik}(\eta_a, \partial_{[b}\eta_{a]}, 0),$$

d. h. S_{ik} ist von $\partial_c g_{ab}$ unabhängig. Auf Grund der Formel (3. 9) beweist das eben den Satz 5.

Die Abhängigkeit von η_a kann nicht eliminiert werden, wie das das Beispiel

$$D_k\eta_i = \partial_{[k}\eta_{i]} - \eta_i\eta_k$$

zeigt.

Nehmen wir jetzt an, daß die kovariante Ableitung $D_k \eta_i$ von η_a unabhängig ist, so besteht der folgende

SATZ 6. *Ist die kovariante Ableitung $D_k \eta_i$ von η_a unabhängig, so ist sie auch von $\partial_c g_{ab}$ unabhängig und $D_k \eta_i$ ist eine Funktion von $\partial_{[b}\eta_{a]}$ und g_{ab}.*

BEWEIS. Durch die Substitution

$$A_a^r = \delta_a^r, \quad A_{a\,c}^{\,r} = -\Gamma_{a\,c}^{\,r}, \quad \eta_a = 0$$

bekommt man aus (3. 6)

$$D_k \eta_i \equiv F_{ik}(\partial_b \eta_a, g_{ab}, \partial_c g_{ab}) = F_{ik}(\partial_b \eta_a, g_{ab}, 0),$$

und das beweist schon die Unabhängigkeit von $\partial_c g_{ab}$. Setzen wir

$$S_{ik}(\partial_b \eta_a, g_{ab}) \stackrel{\text{def}}{=} F_{ik}(\partial_b \eta_a, g_{ab}, 0),$$

so ist

$$D_k \eta_i \equiv S_{ik}(\partial_b \eta_a, g_{ab}). \tag{3. 10}$$

Da die Funktionen S_{ik} einen Tensor bestimmen, bekommt man jetzt statt (3. 6)

$$A_i^r A_k^t S_{rt}(\partial_b \eta_a, g_{ab}) = S_{ik}(A_{a\,b}^{\,r} \eta_r + A_a^r A_b^t \partial_t \eta_r, A_a^r A_b^t g_{rt}).$$

Setzen wir nun wieder

$$A_{a\,b}^{\,r} \eta_r = x_{ab} \equiv x_{ba}, \quad A_a^r = \delta_a^r$$

und nehmen wir dann für x_{ab} die Werte (3. 8), so bekommt man

$$S_{ik}(\partial_b \eta_a, g_{ab}) = S_{ik}(\partial_{[b}\eta_{a]}, g_{ab}),$$

und das beweist nach (3. 10) eben den Satz 7.

Weitere Reduktion der Veränderlichen ist wieder nicht möglich, wie das das Beispiel

$$D_k \eta_i = \partial_{[k}\eta_{r]} g^{rt} \partial_{[t}\eta_{i]}$$

zeigt.

Aus den Sätzen 5 und 6 folgt der

SATZ 7. *Kommen in der Formel der kovarianten Ableitung $D_k \eta_i$ die Größen $\partial_c g_{ab}$ vor, so muß $D_k \eta_i$ auch von g_{ab} und η_a abhängig sein.*

BEWEIS. Wäre nämlich $D_k \eta_i$ von g_{ab} bzw. η_a unabhängig, so würden nach Satz 5 bzw. Satz 6 in der Formel von $D_k \eta_i$ die Größen $\partial_c g_{ab}$ nicht vorkommen können, im Widerspruch zu unserer Annahme.

Es ist ganz offensichtlich, daß die Funktionen S_{ik} in der Formel (3. 2) noch mannigfache, aber nicht ganz beliebige Formen haben können, da ja die kontravarianten g^{ab} durch (1. 1a) allein aus g_{ab} bestimmt werden können; aus η_i, $\nabla_k \eta_i$ und g^{ij} können dann verschiedene rein kovariante Tensoren zweiter Stufe gebildet werden, die kovariante Ableitungen darstellen. Die Funktionen S_{ik} müssen der Gleichung (3. 2a) genügen.

§ 4. Bestimmung der klassischen kovarianten Ableitung

Die in der Differentialgeometrie benützte kovariante Ableitung ist durch (2. 1) bzw. (3. 1) angegeben. Diese kovariante Ableitungen sind in ξ^i, $\partial_k \xi^i$ bzw. in η_i, $\partial_k \eta_i$ homogen und linear, doch charakterisiert diese Eigenschaft diese kovarianten Ableitungen unter den allgemeinsten Typen noch nicht. Das zeigt sich durch die Beispiele:

$$D_k \xi^i = \nabla_k \xi^i + c\delta^i_k \nabla_t \xi^t, \quad D_k \eta_i = \nabla_k \eta_i - c\,\partial_{[k}\eta_{i]} \qquad (c\text{: konst.}),$$

da δ^i_k einen gemischten Tensor darstellt.

Um die kovariante Ableitung (2. 1) bzw. (3. 1) unter der durch die Forderungen A)—C) bestimmten allgemeinen kovarianten Ableitung zu charakterisieren, müssen wir noch weitere Forderungen stellen (vgl. [5], § 7 und § 8).

Um die in den Riemannschen Räumen benützte kovariante Ableitung (2. 1) unter den allgemeinen kovarianten Ableitungen zu charakterisieren, stellen wir die folgenden Forderungen:

1. *$D_k \xi^i$ soll in den $\partial_k \xi^i$ und ξ^i homogen linear sein.*

2. *Ist $\xi^i = 0$ und $\partial_k \xi^i = x^i_k$ in einem beliebigen Punkte $\overset{(0)}{x^i}$, so soll $D_k \xi^i = x^i_k$ sein, wo die x^i_k n^2 beliebige Größen bedeuten.*

3. *Der Koeffizient $\Lambda_i{}^j{}_k$ von ξ^i soll bei der homogen linearen Übertragung in i, k symmetrisch sein.*

Aus dem Satz 2 folgt nach der Forderung 1, daß $D_k \xi^i$ die Form

$$D_k \xi^i = c^{ji}_{kt} \nabla_j \xi^t + c_j{}^i{}_k \xi^j \tag{4. 1}$$

haben muß, wo c^{ji}_{kt} und $c_j{}^i{}_k$ von g_{ab} abhängig sind. Aus der Forderung 2 folgt auf Grund der Formel (2. 1), daß

$$c^{ji}_{kt} x^t_j = x^i_k \tag{4. 2}$$

besteht. Aus (4. 1) und (4. 2) folgt, daß $D_k \xi^i$ die Form

$$D_k \xi^i = \partial_k \xi^i + \Lambda_j{}^i{}_k \xi^j \tag{4. 3}$$

hat, wo nach (2. 1)

$$\Lambda_j{}^i{}_k \overset{\text{def}}{=} \Gamma_j{}^i{}_k + c_j{}^i{}_k(g_{ab}) \tag{4. 4}$$

ist. Auf Grund des Satzes 1 ist aber $c_j{}^i{}_k \equiv 0$; aus (4. 3) und (4 4) folgt dann, daß $D_k \xi^i$ eben die Form (2. 1) hat.

Auf Grund der Forderungen 1—3 ist es also möglich, die in der Riemannschen Geometrie benützte kovariante Ableitung (2. 1) von der allgemeinen, durch die Forderungen A)—C) definierten kovarianten Ableitung zu bestimmen. Jetzt zeigt sich auch die Wichtigkeit des Satzes 1. Der Satz 1

ermöglicht nämlich die explizite Bestimmung der kovarianten Ableitung der Riemannschen Geometrie, ohne die Forderung der Bedingung

$$\nabla_k g_{ij} = 0, \tag{4.5}$$

die sonst bei der Herleitung der Übertragungsparameter $\Gamma_i{}^j{}_k$ üblich ist. (Vgl. [5], § 8, S. 83—89.) (4. 5) drückt aus, daß die kovariante Ableitung metrisch ist. Diese Resultate können wir im folgenden Satz zusammenfassen:

SATZ 8. *Die kovariante Ableitung der Riemannschen Geometrie, die den Forderungen* A)—C) *und* 1—3 *genügt, ist durch* (2. 1) *bestimmt.*

Die Forderungen 1, 2 sind offenbar notwendig, 3 ist aber möglicherweise eliminierbar. Dazu müßte von dem Satz 1 die Bedingung der Symmetrie von $\Gamma_i{}^j{}_k$ in i, k eliminiert werden.

Zum Schluß bemerken wir noch, daß die Forderungen 1—3 die klassische kovariante Ableitung auch für die kovarianten Vektoren charakterisieren, *wenn in* 1—3 *statt* ξ^i *der kovariante Vektor* η_i *gesetzt wird.* Statt (4. 1) kann nach der Forderung 1 für $D_k \eta_i$

$$D_k \eta_i = c_{ik}^{jt} \nabla_j \eta_t - c_i{}^j{}_k \eta_j$$

gesetzt werden, und ebenso wie bei den kontravarianten Vektoren erhält man nach der Forderung 2 die der Gleichung (4. 3) entsprechende Formel

$$D_k \eta_i = \partial_k \eta_i - A_i{}^j{}_k \eta_j, \tag{4.6}$$

wo für $A_i{}^j{}_k$ wieder die Relation (4. 4) gültig ist. Da auch jetzt $c_i{}^j{}_k \equiv 0$ ist, geht (4. 6) auf Grund von (4. 4) in die Formel (3. 1) der klassischen kovarianten Ableitung der kovarianten Vektoren über.

(Eingegangen am 3. November 1959.)

Literaturverzeichnis

[1] J. ACZÉL, Beiträge zur Theorie der geometrischen Objekte. V, *Acta Math. Acad. Sci. Hung.*, **8** (1957), S. 53—64.

[2] S. GOŁĄB, Über den Begriff der kovarianten Ableitung, *Nieuw. Arch. vor Wisk.* (3), **2** (1954), S. 90—96.

[3] A. MOÓR, Über die kovariante Ableitung der Vektoren, *Acta Sci. Math. Szeged*, **19** (1958), S. 237—246.

[4] A. MOÓR, Über Tensoren, die aus angegebenen geometrischen Objekten gebildet sind, *Publ. Math. Debrecen*, **6** (1959), S. 15—25.

[5] J. A. SCHOUTEN und D. J. STRUIK, *Einführung in die neueren Methoden der Differentialgeometrie. I* (Groningen, 1935).

[6] S. GOŁĄB und M. KUCHARZEWSKI, Ein Beitrag zur Komitantentheorie, *Acta Math. Acad. Sci. Hung.*, **11** (1960), S. 173—174.

ÜBER DIE GEOMETRISCHE ANWENDUNG EINES DIFFERENTIALGLEICHUNGSSYSTEMS

Von

L. TAMÁSSY (Debrecen)

(*Vorgelegt von* O. VARGA)

Einleitung

Für die Äquivalenz affinzusammenhängender Punkträume ist das Kriterium wohlbekannt, welches mit Hilfe des Transformationsgesetzes der Krümmungs- und Torsionstensoren, und deren kovarianter Ableitungen ausgedrückt wird.[1] Dieses Kriterium ist lediglich die Bedingung der Lösbarkeit des Differentialgleichungssystems, welches das Transformationsgesetz der Übertragungskoeffizienten ausdrückt. Im § 1 werden wir zeigen, daß die Erfüllung des erwähnten Kriteriums nicht in jedem Falle eine solche Äquivalenz sichert, bei welcher die Punkte der zwei Räume ein-eindeutig einander zugeordnet sind. Der Grund dafür ist, daß über dem Lösungssystem des das Transformationsgesetz der Übertragungskoeffizienten ausdrückenden Differentialgleichungssystems die Umkehrbarkeit nicht in jedem Falle gesichert ist. Solche Räume (für welche wir im § 1 Beispiele geben) werden in der Literatur nicht als äquivalent betrachtet.[2] Im § 2 zeigen wir, falls in der Lösung des Differentialgleichungssystems, welches das Transformationsgesetz der Übertragungskoeffizienten ausdrückt, die Zahl der Parameter genügend groß ist — was ohne Kenntnis der Lösungen leicht zu bestimmen ist[3] — dies schon für die Äquivalenz hinreichend ist. Endlich geben wir im § 3 eine notwendige und hinreichende Bedingung für die Äquivalenz von affinzusammenhängenden Räumen an.

§ 1

Ein wohlbekannter Satz [4] von J. M. THOMAS und O. VEBLEN gibt eine notwendige und hinreichende Bedingung für die Lösbarkeit des gemischten Systems der Form

$$\text{(1)}\quad \text{a)}\quad \frac{\partial \theta^\alpha}{\partial \bar{x}^i} = \psi_i^\alpha(\theta, \bar{x}) \qquad \begin{pmatrix} \alpha = 1, 2, \ldots, M \\ i = 1, 2, \ldots, n \end{pmatrix},$$

$$\text{b)}\quad F_0^{\lambda_0}(\theta, \bar{x}) = 0 \qquad (\lambda_0 = 1, 2, \ldots, S).$$

[1] Siehe z. B. [3], S. 205; [1], S. 76; oder [5], Chap. V, § 9.
[2] Siehe [6], Chap. II, § 1, fünfter Absatz.
[3] Siehe [1], S. 16.

Dieses Kriterium ist bei Untersuchungen der Äquivalenz von differentialgeometrischen Räumen auf die folgende Weise sehr gut zu verwenden. Es seien in zwei verschiedenen Koordinatensystemen zwei gleichartige differentialgeometrische Räume gegeben mit den die Struktur des Raumes bestimmenden geometrischen Objekten und mit den Transformationsgesetzen, nach welchen sich diese geometrischen Objekte bei einer Koordinatentransformation transformieren. Die zwei Räume sind definitionsgemäß dann äquivalent, wenn es eine solche umkehrbare[4] Koordinatentransformation gibt, vermöge der die die Struktur des Raumes bestimmenden geometrischen Objekte ineinander übergehen. Wir können die Transformationsgesetze dieser geometrischen Objekte, oder irgendeine passend umgestaltete Form derselben als ein Differentialgleichungssystem ansehen, in welchem die Komponenten, die in den auf verschiedene Koordinatensysteme bezogenen Räumen gegebenen Objekte als bekannte Funktionen auftreten, und die Unbekannten die transformationsbestimmenden Funktionen sind. Kann man dieses Differentialgleichungssystem auf die Form (1) bringen, so ist die Bedingung der Lösbarkeit mit Hilfe des obigen Kriteriums ausdrückbar. Die Lösbarkeit sichert aber (in sich selbst) von vornherein noch nicht die Umkehrbarkeit der Lösungsfunktionen. In diesem Falle existiert aber die zu der Äquivalenz erforderliche umkehrbare Transformation nicht.

In vielen Fällen — wie mich Prof. O. VARGA aufmerksam machte — folgt die Umkehrbarkeit der die Abbildung beschreibenden Funktionen sofort aus der Lösbarkeit des gemischten Systems. So folgt z. B. bei Räumen, in welchen es einen metrischen Grundtensor zweiter Stufe gibt, aus der Gleichung

$$\bar{g}_{ik} = g_{rs} \frac{\partial x^r}{\partial \bar{x}^i} \frac{\partial x^s}{\partial \bar{x}^k},$$

die zu der Gleichungsgruppe (1, b) gehört, die Gleichung

$$\bar{g} = g \left| \frac{\partial x}{\partial \bar{x}} \right|^2,$$

was wegen der Regularität der aus den Grundtensoren gebildeten quadratischen Matrizen nur dann möglich ist, wenn $\left| \frac{\partial x}{\partial \bar{x}} \right| \neq 0$ ist, d. h. wenn das Funktionensystem $x^i = x^i(\bar{x})$ umkehrbar ist.

Um zu beweisen, daß aber die in der Lösung vorkommenden, die Abbildung des einen Koordinatensystems auf das andere beschreibenden Funktionen nicht notwendig ein umkehrbares System bilden, betrachten wir das folgende einfache Beispiel. Wir nehmen zwei zweidimensionale affin-

[4] Siehe [6], Chap. II, § 1.

zusammenhängende Punkträume. Der erste sei auf das Koordinatensystem (x, y), der andere auf das Koordinatensystem $(\bar{x}, \bar{y})$ bezogen. Wir betrachten zwischen den zwei Räumen die Abbildung

$$x = \bar{x} + \bar{y}, \quad y = \bar{x} + \bar{y}. \tag{2}$$

Diese Zuordnung bildet die Punkte des Raumes $(\bar{x}, \bar{y})$ auf die Gerade

$$y = x \tag{3}$$

des Raumes (x, y) ab. Das Transformationsgesetz der Übertragungskoeffizienten hat die Form

$$\bar{\Gamma}_{b\ c}^{\ a} = \bar{\Gamma}_{j\ k}^{\ i} \frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial x^k}{\partial \bar{x}^c} \frac{\partial \bar{x}^a}{\partial x^i} + \frac{\partial^2 x^i}{\partial \bar{x}^b \partial \bar{x}^c} \frac{\partial \bar{x}^a}{\partial x^i} \tag{4}$$

oder

$$\Gamma_{j\ k}^{\ i}(x) \frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial x^k}{\partial \bar{x}^c} = \bar{\Gamma}_{b\ c}^{\ a}(\bar{x}) \frac{\partial x^i}{\partial \bar{x}^a} - \frac{\partial^2 x^i}{\partial \bar{x}^b \partial \bar{x}^c}. \tag{5}$$

Geben wir jetzt die Γ so an, daß sie längs (3) Konstanten sind, und dort die Bedingungen $\Gamma_{j\ k}^{\ i} = \Gamma_{k\ j}^{\ i} = \Gamma_{k\ j}^{\ l}$ erfüllen, während sie an anderen Stellen asymmetrisch sind. Man kann leicht sehen, daß (5) längs (3) für die $\bar{\Gamma}$ eine Lösung besitzt. Wir wählen die $\bar{\Gamma}$ im ganzen Raum gleich diesen Konstanten. So bestimmen wir zwei affinzusammenhängende Räume, die nicht äquivalent sind, weil in dem ersten Raum das Objekt des Zusammenhanges symmetrisch ist, während der andere Raum kein zweidimensionales Gebiet besitzt, wo das Objekt des Zusammenhanges symmetrisch wäre. Das Differentialgleichungssystem (5) ist jedoch lösbar, weil ja die Funktionen (2) bei den oben beschriebenen Γ und $\bar{\Gamma}$ das Differentialgleichungssystem (5) in eine Identität in $\bar{x}$ und $\bar{y}$ verwandeln. Im § 2 werden wir ein Beispiel erörtern, das gleichzeitig auch ein Beispiel für den Fall bedeutet, wo die beiden Räume ohne Torsion sind, der erste überall verschwindende, der zweite überall nichtverschwindende Krümmung hat, und das Differentialgleichungssystem (5) doch lösbar ist.

So gibt die Erfüllung des Kriteriums von THOMAS und VEBLEN im Falle des auf die Gestalt (1) gebrachten Differentialgleichungssystems (5) nur eine notwendige, aber in sich selbst nicht immer hinreichende Bedingung für die Äquivalenz.

Wir bemerken, daß das ursprüngliche Transformationsgesetz der Γ (4) ist. Dies ist aber bei gegebenem Γ und $\bar{\Gamma}$ für das Funktionensystem $x^i = x^i(\bar{x})$ kein Differentialgleichungssystem, weil in dies nicht nur die unbekannten Funktionen und deren partielle Ableitungen, sondern auch die partiellen Ableitungen der Inversen des gesuchten Funktionensystems eingehen. So ist der

Übergang auf die Form (5) notwendig. (4) ist wirklich mit (5) bezüglich der umkehrbaren Transformationen äquivalent, weil in einem solchen Falle aus (4) (5) folgt, und umgekehrt. Jetzt wissen wir aber noch nicht, ob es Funktionen gibt, die eine solche Transformation zustande bringen. Wir müssen daher bei dem Übergang die Existenz solcher Funktionen (stillschweigend) voraussetzen, genauer, der Übergang ist nur dann erlaubt, wenn die Existenz solcher Funktionen mindestens nachträglich bewiesen wird.

§ 2

Nunmehr wollen wir untersuchen, wie man die mit Hilfe des Thomas—Veblenschen Kriteriums beschriebene notwendige Bedingung der Äquivalenz im Falle n-dimensionaler affinzusammenhängender Punkträume mit hinreichenden Bedingungen ergänzen kann.

Das auf die Form (§ 1, (1)) gebrachte Differentialgleichungssystem (§ 1, (5)) hat die Gestalt

$$\begin{aligned} &\frac{\partial x^i}{\partial \bar{x}^a} = u^i_a(\bar{x}), \\ (1)\qquad & \\ &\frac{\partial u^i_b}{\partial \bar{x}^c} = \Gamma_j{}^i{}_k(x)\, u^j_b(\bar{x})\, u^k_c(\bar{x}) - \bar{\Gamma}_b{}^a{}_c(\bar{x})\, u^i_a(\bar{x}). \end{aligned}$$

Die Lösung desselben hat die Form

$$(2)\quad \begin{aligned} &\text{a)} \qquad x^i = x^i(\bar{x}, c), \\ &\text{b)} \qquad u^i_a = u^i_a(\bar{x}, c), \end{aligned}$$

wo c die in der Lösung auftretenden Parameter bezeichnet. Eine hinreichende Bedingung der Äquivalenz ist neben der Lösbarkeit von (1), daß das die Transformation beschreibende Funktionensystem (2, a) in jedem Punkt $\bar{x}_0$ mindestens für ein Wertsystem c_0 umkehrbar ist, d. h. die Determinante

$$(3)\qquad \left| u^i_a(\bar{x}_0, c) \right|$$

in c nicht identisch verschwindet. Ist die Lösung (2) schon bekannt, so ist es sehr leicht zu entscheiden, ob ein solches Wertsystem für die Parameter existiert oder nicht. Wir möchten aber eine solche Bedingung finden, welche die Äquivalenz ohne auch Kenntnis der Lösung von (1) sichert.

Wir erhalten die Lösung der in (1) vorkommenden n^2+n unbekannten Funktionen folgendermaßen:[5] Wir bilden die Integrabilitätsbedingungen von (1), dann differenzieren wir diese nach $\bar{x}$, und wir substituieren die Werte

[5] Siehe [2], § 1.

von (1) in den gewonnenen Ausdruck. Wir wiederholen diesen Prozeß, bis die mit diesem Verfahren gebildeten neuen Ausdrücke mit Rücksicht auf die Vorangehenden lauter Identitäten geben. Unter diesen Gleichungen gibt es p unabhängige. Aus diesen Gleichungen kann man nun die übrigen unbekannten Funktionen mit diesen p Funktionen ausdrücken. Die Lösungen für diese letzteren p unbekannten Funktionen erhalten wir aus einem vollständig integrablen Differentialgleichungssystem, z. B. in der Form einer Potenzreihe in der Umgebung von $\bar{x}_0$. So enthält jede dieser Funktionen eine frei wählbare Konstante, die der Wert der Lösungsfunktion im Punkte $\bar{x}_0$ ist. So enthalten diese Funktionen insgesamt p Konstanten $c^1, c^2, \ldots, c^p$. Diese Funktionen sollen weiterhin „frei wählbar" genannt werden. Die weitere n^2+n-p Funktionen sind durch die schon gelösten Funktionen ausgedrückt. Man kann erwarten, daß im Falle einer genügend großen Anzahl der Parameter sich immer ein solches c_0 finden läßt, für welches (3) nicht verschwindet.

Wir zeigen, daß es für $p=n^2+n-(n-1)=n^2+1$ noch immer ein solches c_0 gibt, für welches (3) nicht verschwindet. — Gemäß unserer Bedingung ist die Nummer der Parameter, und so die Anzahl der frei wählbaren Funktionen höchstens um $n-1$ kleiner als die Gesamtanzahl der Funktionen. So kommen unter den Elementen der Determinante (3) höchstens $n-1$ solche vor, die nicht frei wählbare Funktionen sind. Diese Determinante besitzt also mindestens eine Reihe, welche aus lauter frei wählbaren Funktionen besteht. Wir wählen eine solche Reihe. Wir wählen die Konstanten, also die in $\bar{x}_0$ angenommenen Werte sämtlicher in dieser Reihe stehenden Funktionen gleich Null, mit Ausnahme eines einzigen Funktionenwertes, in deren Kolumne auch eine von den frei wählbaren verschiedene Funktion steht. Der Wert dieser Funktion in $\bar{x}_0$ sei gleich 1. Entwickeln wir die Determinante (3) nach dieser Reihe, so erhalten wir eine einzige Determinante vom Rang $n-1$, die mit (3) bis auf das Vorzeichen gleichen Wert besitzt. In dieser Determinante vom Rang $n-1$ befinden sich höchstens $n-2$ solche Elemente, die nicht frei wählbare Funktionen sind, weil es in (3) höchstens $n-1$ solche gab, und weil sich in der weggelassenen Kolumne mindestens eine solche befindet. So ist diese Determinante vom ähnlichen Typus wie die vorangehende, so daß sich das Vehrfahren wiederholen läßt. Wir gelangen mit dem letzten Schritt dieses Verfahrens offensichtlich zu einer Determinante vom Rang 1, deren einziges Element eine frei wählbare Funktion ist. Also läßt sich für den obenerwähnten Wert von p tatsächlich ein solches c_0 finden, für welches (3) nicht verschwindet. So erhalten wir den folgenden

SATZ 1. *Für die Äquivalenz zweier affinzusammenhängender Räume ist es hinreichend, wenn das Differentialgleichungssystem* (1) *lösbar ist, und in der Lösung mehr als n^2 unabhängige Parameter vorkommen.*

Ist p nicht größer als n^2, z. B. $p=n^2$, so ist in (3) lediglich die Existenz von n^2-n frei wählbarer Funktionen gesichert. Falls die übrigbleibenden n Funktionen von (3) sich in derselben Zeile befinden, und diese Zeile eine lineare Kombination der Übrigen ist, so verschwindet (3) identisch.

Das dies wirklich vorkommen kann, beweisen wir mit einem neuen Beispiel. — Die Übertragungskoeffizienten des zweidimensionalen Raumes (x^1, x^2) seien $\Gamma_j{}^i{}_k(x)\equiv 1$, und diejenigen des Raumes $(\bar{x}^1, \bar{x}^2)$

$$\bar{\Gamma}_1{}^1{}_1(\bar{x})\equiv 2\alpha, \quad \bar{\Gamma}_1{}^2{}_1(\bar{x})\equiv 0, \quad \bar{\Gamma}_1{}^1{}_2(\bar{x})\equiv \beta, \quad \bar{\Gamma}_1{}^2{}_1(\bar{x})\equiv \alpha,$$

$$\bar{\Gamma}_2{}^1{}_1(\bar{x})\equiv \beta, \quad \bar{\Gamma}_2{}^2{}_1(\bar{x})\equiv \alpha, \quad \bar{\Gamma}_2{}^1{}_2(\bar{x})\equiv 0, \quad \bar{\Gamma}_2{}^2{}_2(\bar{x})\equiv 2\beta$$

(wo α und β nichtverschwindende Konstanten sind). Bilden wir mit diesen Γ und $\bar{\Gamma}$ die Integrabilitätsbedingungen von (1), welche durch das Transformationsgesetz des Krümmungstensors ausgedrückt werden, so finden wir mit Rücksicht auf das Verschwinden der Krümmung des Raumes x, daß es unter diesen Gleichungen zwei unabhängige gibt, welche auch verträglich sind:

$$\beta\frac{\partial x^l}{\partial \bar{x}^1}-\alpha\frac{\partial x^l}{\partial \bar{x}^2}=0 \qquad (l=1,2). \tag{4}$$

Differenzieren wir diese Gleichungen nach $\bar{x}^j$, so erhalten wir mit Rücksicht auf (1) und (4) Identitäten. Daher ist das Differentialgleichungssystem (1) lösbar. Die Zahl der unabhängigen Gleichungen ist 2, daher ist die Zahl der in der Lösung vorkommenden Parameter $6-2=4$, was um eins kleiner als n^2+1 ist. Die diesem Falle entsprechenden Funktionen (2, a) können nirgends und für keinen Wert der vier Parameter ein umkehrbares System bilden, d. h. die Determinante (3) muß für jeden Wert von c und $\bar{x}$ verschwinden. Sonst wären nämlich die zwei Räume oder einige Teile derselben äquivalent, was in Widerspruch zu der Tatsache steht, daß die Krümmung des Raumes x Null ist, während die Krümmung des Raumes $\bar{x}$ nicht verschwindet.[6] Daher ist zwar zur Äquivalenz im allgemeinen keine für die Anzahl der Parameter auferlegte Bedingung notwendig, doch gilt der folgende

SATZ 2. *Die im vorangehenden Satz festgestellte Anzahl n^2 stellt den niedrigsten Wert dar, für welche die Äquivalenz zusammen mit der Lösbarkeit von* (1) *immer (d. h. für jede Dimensionszahl n) gesichert ist.*

Es ist wahrscheinlich, daß n_0^2 auch für jede spezielle Wahl $n=n_0$ der kleinste Wert ist.

[6] Dies bedeutet gleichzeitig ein Beispiel für solche, in § 1 erwähnte Räume, welche ohne Torsion sind, wobei die Krümmung der ersten Raumes überall, diejenige des zweiten nirgends verschwindet, und das Differentialgleichungssystem (1) dennoch lösbar ist.

§ 3

SATZ 3. *Für die Äquivalenz der affinzusammenhängenden Räume $\Gamma(x)$ und $\overline{\Gamma}(\bar{x})$ ist es notwendig und hinreichend, daß das aus den Differentialgleichungen*

$$(1)\quad \begin{aligned} &\text{a, I)} && \frac{\partial x^i}{\partial \bar{x}^b} = p^i_b(\bar{x}), \\ &\text{b, I)} && \frac{\partial z^a}{\partial y^i} = q^a_i(y), \\ &\text{a, II)} && \frac{\partial p^i_b}{\partial \bar{x}^c} = \overline{\Gamma}_b{}^a{}_c(\bar{x}) p^i_a(\bar{x}) - \Gamma_l{}^i{}_k(x) p^l_b(\bar{x}) p^k_c(\bar{x}), \\ &\text{b, II)} && \frac{\partial q^a_d}{\partial y^c} = -\overline{\Gamma}_b{}^a{}_f(z) q^f_c(y) q^b_d(y) + \Gamma_d{}^i{}_c(y) q^a_i(y) \end{aligned}$$

und aus den skalaren Gleichungen

$$(2)\qquad x^i(z(y)) = y^i$$

bestehende gemischte System eine Lösung besitze.

Zunächst ist Satz 3 hinreichend. Existiert nämlich ein (1) und (2) befriedigendes Funktionensystem

$$(3)\qquad x(\bar{x}),\quad z(y),\quad p(\bar{x}),\quad q(y),$$

so ergibt sich durch Differentiation von (2) $\dfrac{\partial x^i}{\partial \bar{x}^r}\dfrac{\partial z^r}{\partial y^j} = \delta^i_j$, d. h.

$$(4)\qquad p^i_r(z(y)) q^r_j(y) = \delta^i_j .$$

Daher ist

$$|p^i_r|\,|q^r_j| = 1,$$

d. h.

$$|p^i_b| \neq 0,$$

und dies bedeutet, daß die Lösung

$$(5)\qquad x^i = x^i(\bar{x})$$

umkehrbar ist. Dies ist wegen (1, a) eine Lösung von (§ 1, (5)), und wegen der Umkehrbarkeit auch von (§ 1, (4)).

Wir zeigen, daß die Bedingung des Satzes auch notwendig ist, d. h. wenn $\Gamma(x)$ und $\overline{\Gamma}(\bar{x})$ äquivalent sind, so existiert ein (1) und (2) befriedigendes Funktionensystem (3). Seien nämlich $\Gamma(x)$ und $\overline{\Gamma}(\bar{x})$ äquivalent. Dann

existiert nach der Bedeutung der Äquivalenz das Funktionensystem (5), welches (§ 1, (4)), (§ 1, (5)), und daher auch (1, a) befriedigt. (1, a) muß also in diesem Falle gültig sein. Seien die z^i die mit $\bar{x}^i = \bar{x}^i(x)$ bezeichneten Inversen der jetzt nach unserer Annahme existierenden umkehrbaren Funktionen (5) (die Veränderlichen der z^i bezeichnen wir mit y), und die q^i_j seien ihre partiellen Derivierten. Daher besteht (2) offensichtlich identisch. Da hier die x und z zueinander invers sind, so besteht (4), und auch

$$q^d_p(x(\bar{x}))p^p_s(\bar{x}) = \delta^d_s. \tag{6}$$

Daher ergibt sich aber (1, b, II) aus (4) durch partielle Ableitung mit Rücksicht auf (1, a, II) und (6). Es muß daher in diesem Falle auch (1, b) bestehen.

(1) sichert lediglich, daß nicht nur ein solches Funktionensystem existieren muß, welches den Raum $\bar{x}$ auf dem Raum x abbildet (vorläufig nicht ein-eindeutigerweise, sondern vielleicht nur eindeutigerweise), und welches natürlich die $\bar{\Gamma}(\bar{x})$ nach dem gegebenen Gesetz in die $\Gamma(x)$ der entsprechenden Punkte x transformiert — dies ist durch (1, a) gesichert —; sondern auch ein Funktionensystem (welches vorläufig wieder nicht notwendigerweise ein-eindeutig, sondern nur eindeutig sein könnte), welches eine Abbildung in der umgekehrten Richtung mit ähnlichen Eigenschaften zustande bringt — dies ist durch (1, b) gesichert. Wegen (2) müssen aber diese Funktionensysteme zueinander invers sein, d. h. es müssen die beiden Funktionensysteme dieselbe ein-eindeutige Abbildung zustande bringen.

Endlich drücken wir die Bedingung der Äquivalenz mit Hilfe der Krümmungs- und Torsionstensoren aus. Wir wenden auf das gemischte System (1), (2) das im § 1 erwähnte Kriterium von THOMAS und VEBLEN an. Hier spielen die Funktionen (3) die Rolle des θ von (§ 1, (1)) und an Stelle von $\bar{x}$ treten die $\bar{x}$ und y. Wir bilden die Integrabilitätsbedingungen von (1) und die partiellen Ableitungen von (2). Für die Integrabilitätsbedingungen von (1, a, I) ergibt sich mit Hilfe von (1, a, II)

$$\bar{S}_b{}^a{}_c(\bar{x})p^i_a(\bar{x}) = S_j{}^i{}_k(x)p^j_b(\bar{x})p^k_c(\bar{x}), \tag{7, a, I}$$

wo S der Torsionstensor ist. Die Integrabilitätsbedingungen von (1, b, I) sind mit Rücksicht auf (1, b, II)

$$\bar{S}_b{}^a{}_f(z)q^f_c(y)q^b_d(y) = S_d{}^i{}_c(y)q^a_i(y). \tag{7, b, I}$$

Diejenigen von (1, a, II) und (1, b, II) sind ähnlicherweise

$$\bar{R}^a{}_{bcd}(\bar{x})p^i_a(\bar{x}) = R^i{}_{jkl}(x)p^j_b(\bar{x})p^k_c(\bar{x})p^l_d(\bar{x}), \tag{7, a, II}$$

und

$$\bar{R}^a{}_{bcd}(z)q^b_j(y)q^c_k(y)q^d_l(y) = R^i{}_{jkl}(y)q^a_i(y). \tag{7, b, II}$$

Die weiteren Ableitungen drücken mit Rücksicht auf (1) und (7) das tensorielle Transformationsgesetz der kovarianten Derivierten der Krümmungs- und Torsionstensoren in dualer Weise aus, ähnlich wie es bei (7) geschah.

Die Ableitungen von (2) sind (4), die für p und q, als für selbständige unbekannte Funktionen neue Relationen sind. Die Derivierte von (4) gibt aber schon keine neue Relation, sondern reduziert sich auf (1, b, II). Aus (2) folgt nämlich

$$z^d(x(z(y))) = z^d(y), \tag{8}$$

und durch Ableitung von (8) nach $z^\nu(y)$ erhalten wir gemäß (1, I) und (2)

$$q^d_p(y)\, p^p_\nu(z(y)) = \delta^d_\nu. \tag{9}$$

Daher reduziert sich die Ableitung von (4) mit Hilfe von (9) und (1, a, II) auf (1, b, II). Wir bezeichnen die Gleichungen, die das tensorielle Transformationsgesetz der k-ten kovarianten Ableitungen der Krümmungs- und Torsionstensoren in einer zu (7) ähnlichen dualen Form ausdrücken, kurz mit $F^\lambda_k = 0$ $(\lambda = 1, 2, \ldots, N_k)$. Dann gilt der

Satz 4. *Für die Äquivalenz der Räume $\Gamma(x)$ und $\bar{\Gamma}(\bar{x})$ ist die Existenz einer natürlichen Zahl N notwendig und hinreichend, für welche die Gleichungen* (2), (4), (7), $F^\mu_1 = 0$, $F^\nu_2 = 0, \ldots, F^\lambda_N = 0$ *ein verträgliches System bilden, und für welche die* $F^\varkappa_{N+1} = 0$ *Folgerungen der vorangehenden Gleichungen sind.*

(Eingegangen am 5. November 1959.)

Literaturverzeichnis

[1] L. P. Eisenhart, *Non-Riemannian geometry* (New York, 1927).
[2] L. P. Eisenhart, *Continuous groups of transformations* (Princeton, 1933).
[3] T. Y. Thomas, *The differential invariants of generalized spaces* (Cambridge, 1934).
[4] J. M. Thomas and O. Veblen, Projective invariants of affine geometry of paths, *Annals of Math.*, **27** (1925—26), S. 279—296.
[5] O. Veblen, *Invariants of quadratic differential forms* (Cambridge, 1927).
[6] O. Veblen and J. Whitehead, *The foundations of differential geometry*, Cambridge Tracts, No. 29 (1953).

ÜBER DIE ZERLEGBARKEIT VON FINSLERSCHEN RÄUMEN

Von

O. VARGA (Budapest), korrespondierendem Mitglied der Akademie

Für die Riemannsche Geometrie sind die Kriterien für die Darstellung eines n-dimensionalen Riemannschen Raumes als Produkt zweier r- bzw. $n-r$-dimensionaler Riemannscher Räume wohl bekannt.[1] Beschränkt man sich auf lokale Untersuchungen, dann ist die erwähnte Zerfällung gleichbedeutend a) mit der Existenz eines unbeschränkt parallelverschiebbaren r-dimensionalen Vektorraumes, b) mit der Existenz eines kovarianten symmetrischen positiv semidefiniten Tensors zweiter Stufe, dessen invariantes Differential identisch verschwindet, c) mit der Existenz einer reduziblen Holonomiegruppe.

Aus unseren Darlegungen wird hervorgehen, daß in der Finslerschen Geometrie die Existenz unbeschränkt parallelverschiebbarer r-dimensionaler Vektorräume V_r keineswegs die Möglichkeit der Zerlegbarkeit nach sich zieht. Es wird sich herausstellen, daß für die Existenz von absolut parallelverschiebbaren Vektorräumen V_r, das Vorhandensein eines kovarianten positiv semidefiniten symmetrischen Tensors zweiter Stufe mit identisch verschwindendem invariantem Differential eine notwendige und hinreichende Bedingung ist, zu der noch zwei weitere Bedingungen hinzukommen, die dann zusammen notwendig und hinreichend für die Produktdarstellung des F_n sind.

Die Reduzibilität der einem Linienelement angehörenden Holonomiegruppe ist, wie im Falle der Riemannschen Geometrie, äquivalent mit der Existenz von absolut parallelverschiebbaren Vektorräumen, falls die Länge eines Kurvenbogens nur von dem Bogen selbst, hingegen nicht von seiner Orientierung abhängt.

Unsere Ausführungen sind von lokalem Charakter. Wir setzen stets die Existenz eines lokalen Koordinatensystems (x^i, v^j) voraus, für welches sämtliche auftretende Größen hinreichend oft differenzierbar sind.

Wir beginnen mit folgender

Definition. *Der n-dimensionale Finslersche Raum F_n mit der Grundfunktion $L(x, v)$ werde als Produktraum eines r- bzw. $(n-r)$-dimensionalen F_r*

[1] Vgl. J. A. Schouten, *Ricci-Calculus*, 2 Auflage (Berlin—Göttingen—Heidelberg, 1954), insbes. S. 285—287, wo auch weitere Literaturangaben zu finden sind.

bzw. F_{n-r} *betrachtet, falls es zwei Grundfunktionen* $L_1(x^1, \ldots, x^r; v^1, \ldots, v^r)$ *bzw.* $L_2(x^{r+1}, \ldots, x^n; v^{r+1}, \ldots, v^n)$ *gibt, für die*

$$L^2 = L_1^2 + L_2^2$$

gilt, und

$$L_1(x^{\varrho'}, 0) \equiv 0, \quad L_2(x^{\varrho''}, 0) \equiv 0$$

besteht.

Der Indexbereich der Zahlen von 1 bis r sei mit griechischen Buchstaben und einem darüber gesetzten Strich, derjenige der Zahlen von $r+1$ bis n mit griechischen Buchstaben mit zwei darüber gesetzten Strichen bezeichnet.

Wie in der Riemannschen Geometrie, zerfallen die Komponenten irgendeiner Größe des F_n in zwei Gruppen, nämlich Komponenten mit einmal und Komponenten mit zweimal gestrichenen Indizes. Natürlich werden dabei nur solche Koordinatentransformationen zugelassen, die die einmal bzw. zweimal gestrichenen Koordinaten getrennt ineinander überführen.

Ein orthonormiertes n-Bein von Vektoren bestehe aus r Vektoren des F_r und $n-r$ Vektoren des F_{n-r}. Werden dieselben mit $\underset{(\varrho')}{\xi}{}^{\varkappa'}$ bzw. $\underset{(\varrho'')}{\xi}{}^{\varkappa''}$ bezeichnet, so gilt für die beiden Komponentengruppen des metrischen Fundamentaltensors

$$g_{\varrho'\varkappa'} = \underset{(\mu')}{\xi}{}_{\varrho'} \underset{(\mu')}{\xi}{}_{\varkappa'}, \quad g_{\varrho''\varkappa''} = \underset{(\mu'')}{\xi}{}_{\varrho''} \underset{(\mu'')}{\xi}{}_{\varkappa''}.$$

In F_n bestimmt der Tensor p'_{ik}, für den

$$p'_{\varrho'\varkappa'} = g_{\varrho'\varkappa'}, \quad p'_{\varrho''\varkappa''} = 0$$

gilt, einen r-dimensionalen Vektorraum V_r, und entsprechend definiert der Tensor p''_{ik}, für den

$$p''_{\varrho'\varkappa'} = 0, \quad p''_{\varrho''\varkappa''} = g_{\varrho''\varkappa''}$$

gilt, einen $(n-r)$-dimensionalen Vektorraum V_{n-r}, der zu V_r im Sinne der Raummetrik orthogonal ist. Wegen des Ricci-Lemmas ist

$$Dp'_{ik} = 0, \tag{1}$$

$$Dp''_{ik} = 0. \tag{1a}$$

Betrachten wir zwei solche Vektorfelder $\xi^{\varrho'}$ und $\eta^{\varrho'}$ des F_n, die Vektorfelder des F_r sind, dann wird das Feld $\zeta^{\varrho'}$, das durch Bildung des Kommutators

$$\zeta^{\varrho'} = [\xi\eta]^{\varrho'} = \frac{\partial \xi^{\varrho'}}{\partial x^{\sigma'}} \eta^{\sigma'} - \frac{\partial \eta^{\varrho'}}{\partial x^{\sigma'}} \xi^{\sigma'}$$

entsteht, ebenfalls ein Feld des V_r sein. Entsprechendes gilt für den Kommutator zweier Vektorfelder des F_{n-r}. Wir werden diesen Tatbestand als Abgeschlossenheit der Kommutatorbildung hinsichtlich V_r bzw. V_{n-r} bezeichnen.

Insbesondere gilt also

$$[\underset{(\varrho')}{\xi}\ \underset{(\sigma')}{\xi}]^{\varkappa'} = c^{\mu'}_{\varrho'\sigma'}\, \underset{(\mu')}{\xi}{}^{\varkappa'}, \tag{2}$$

$$[\underset{(\varrho'')}{\xi}\ \underset{(\sigma'')}{\xi}]^{\varkappa''} = c^{\mu''}_{\varrho''\sigma''}\, \underset{(\mu'')}{\xi}{}^{\varkappa''}. \tag{3}$$

Aus (1) und (2) folgt, daß a) die Existenz eines positiv semidefiniten Tensors zweiter Stufe p'_{ik} vom Range r mit verschwindendem invariantem Differential, b) die Abgeschlossenheit der Kommutatorbildung hinsichtlich des durch p'_{ik} bestimmten V_r sowie seines Normalraumes V_{n-r} notwendige Bedingungen für die Produktdarstellung des F_n sind.

Wir zeigen nun, daß diese Bedingungen auch hinreichend sind. Es sei also ein F_n gegeben, für den p'_{ik} ein symmetrisches, positiv semidefinites Tensorfeld vom Rang r sei, für das (1) gilt.

Sind $\underset{(\varrho')}{\xi}{}^{i}$ r orthonormierte Vektoren, die den durch p'_{ik} bestimmten V_r aufspannen, so daß also

$$p'_{ik} = \underset{(\varrho')}{\xi}{}_{i}\, \underset{(\varrho')}{\xi}{}_{k}$$

besteht, so möge für diese Vektoren die Abgeschlossenheitsbedingung (2) gelten:

$$[\underset{(\varrho')}{\xi}\ \underset{(\mu')}{\xi}]^{i} = c^{\sigma'}_{\varrho'\mu'}\, \underset{(\sigma')}{\xi}{}^{i}. \tag{2'}$$

Der Normalraum V_{n-r} von V_r werde durch die Vektoren $\underset{(\varrho'')}{\xi}{}^{i}$ aufgespannt, die zusammen mit den $\underset{(\varrho')}{\xi}{}^{i}$ ein orthonormiertes r-Bein bilden. Wir fordern für die Vektoren des V_{n-r} die Abgeschlossenheitsbedingung (3), d. h.

$$[\underset{(\varrho'')}{\xi}\ \underset{(\mu'')}{\xi}]^{i} = c^{\sigma''}_{\varrho''\mu''}\, \underset{(\sigma'')}{\xi}{}^{i}. \tag{3'}$$

Da

$$g_{ik} = \underset{(\varrho')}{\xi}{}_{i}\, \underset{(\varrho')}{\xi}{}_{k} + \underset{(\varrho'')}{\xi}{}_{i}\, \underset{(\varrho')}{\xi}{}_{k} = p'_{ik} + p''_{ik}$$

ist, folgt aus (1), daß auch

$$D p''_{ik} = 0 \tag{4}$$

ist.

Aus (2′) und (3′) folgt, daß

$$\frac{\partial F}{\partial x^{i}}\, \underset{(\varrho')}{\xi}{}^{i} = 0, \tag{5}$$

$$\frac{\partial F}{\partial x^{i}}\, \underset{(\varrho'')}{\xi}{}^{i} = 0 \tag{6}$$

vollständige Systeme bilden. (6) hat daher r unabhängige Lösungen $\bar{x}^1, \ldots, \bar{x}^r$, und die unabhängigen Lösungen von (5) mögen mit $\bar{x}^{r+1}, \ldots, \bar{x}^n$ bezeichnet werden.

Da für die $\bar{x}^1, \ldots, \bar{x}^n$

$$\frac{\partial(\bar{x})}{\partial(x)} \neq 0$$

gilt, können wir die $\bar{x}^i$ als neue Veränderliche in F_n einführen. Wegen (5) und (6) erhält man dann für $\underset{(k)}{\bar{\xi}^i}$

$$\underset{(\varrho')}{\bar{\xi}^{\varrho''}} = \underset{(\varrho'')}{\bar{\xi}^{\varrho'}} = 0. \tag{7}$$

Hieraus folgt sofort, daß

$$\bar{g}^{\varrho'\varrho''} = \underset{(\mu')}{\bar{\xi}^{\varrho'}} \underset{(\mu')}{\bar{\xi}^{\varrho''}} + \underset{(\mu'')}{\bar{\xi}^{\varrho'}} \underset{(\mu'')}{\bar{\xi}^{\varrho''}} = 0, \tag{8}$$

und hieraus weiter

$$\bar{g}_{\varrho'\varrho''} = 0. \tag{9}$$

Aus (1) folgt für die Vektoren $\underset{(\varrho')}{\xi^i}$ von V_r

$$D \underset{(\varrho')}{\xi^i} = c_{\varrho'\sigma'} \underset{(\sigma')}{\xi^i}, \tag{10}$$

und entsprechend

$$D \underset{(\varrho'')}{\xi^i} = c_{\varrho''\sigma''} \underset{(\sigma'')}{\xi^i}. \tag{10'}$$

Auf Grund von (1) kommt, wenn man die durch

$$D \underset{(\varrho')}{\xi^i} = \underset{(\varrho')}{\xi^i}_{;k} \omega^k + \underset{(\varrho')}{\xi^i}_{|k} dx^k$$

definierten beiden kovarianten Ableitungen einführt,[2]

$$\underset{(\varrho')}{\xi^i}_{;k} = h^{\sigma'}_{k\varrho'} \underset{(\sigma')}{\xi^i}, \tag{11}$$

$$\underset{(\varrho')}{\xi^i}_{|k} = r^{\sigma'}_{k\varrho'} \underset{(\sigma')}{\xi^i}. \tag{12}$$

Beachtet man den expliziten Ausdruck

$$\underset{(\varrho')}{\xi^i}_{;k} = L \frac{\partial \underset{(\varrho')}{\xi^i}}{\partial \dot{x}^k} + A^i_{km} \underset{(\varrho')}{\xi^m} \tag{13}$$

der linken Seite von (11) und geht zu den Koordinaten $\bar{x}^i$ über und wählt

[2] Vgl. E. Cartan, *Les espaces de Finsler*, Actualités scientifiques et industrielles (Paris, 1934), S. 17—19.

den freien Index i aus dem Indexgebiet ϱ'', so kommt wegen (11) und (13)

$$\bar{A}^{\varrho''}_{\varrho'\mu'}=0, \tag{14'}$$

und daher

$$\bar{A}_{\varrho'\mu'\varrho''}=0. \tag{14}$$

Entsprechend findet man

$$\bar{A}_{\varrho''\mu''\varrho'}=0. \tag{15}$$

Die Gleichungen (14) und (15) besagen, daß $\bar{g}_{\varrho'\mu'}$ nur von $\bar{v}^1,\ldots,\bar{v}^r$ und entsprechend $\bar{g}_{\varrho''\mu''}$ nur von $\bar{v}^{r+1},\ldots,\bar{v}^n$ abhängen.

Demnach gilt

$$\begin{aligned}\bar{L}^2(\bar{x}^1,\ldots,\bar{x}^n,\ \bar{v}^1,\ldots,\bar{v}^n)&=L_1^2(\bar{x}^1,\ldots,\bar{x}^n,\ \bar{v}^1,\ldots,\bar{v}^r)+\\&+L_2^2(\bar{x}^1,\ldots,\bar{x}^n,\ \bar{v}^{r+1},\ldots,\bar{v}^n).\end{aligned} \tag{a}$$

Aus den Gleichungen (12) erhält man wegen

$$\underset{(\varrho')}{\bar{\xi}^i}{}_{|k}=\frac{\partial \underset{(\varrho')}{\bar{\xi}^i}}{\partial x^k}-\frac{\partial \underset{(\varrho')}{\bar{\xi}^i}}{\partial v^s}\,\bar{G}^s_k+\bar{\Gamma}^{*i}_{km}\underset{(\varrho')}{\bar{\xi}^m}$$

und (7)

$$0=\underset{(\varrho')}{\bar{\xi}^{\varrho''}}{}_{|\varkappa'}=\bar{\Gamma}^{*\varrho''}_{\varkappa'\mu'}\underset{(\varrho')}{\bar{\xi}^{\mu'}},$$

d. h.

$$\bar{\Gamma}^{*\varrho''}_{\varkappa'\mu'}=0, \tag{16}$$

und falls statt μ' das Indexgebiet μ'' gewählt wird,

$$\bar{\Gamma}^{*\varrho'}_{\varkappa'\mu''}=0. \tag{16'}$$

Hieraus ergibt sich aber

$$\bar{G}^{\varrho'}_{\varrho''}=0, \tag{17}$$

und auf entsprechende Weise schließt man, daß

$$\bar{G}^{\varrho''}_{\varrho'}=0 \tag{17'}$$

ist. Die Gleichungen (16) uud (16') besagen, daß

$$\bar{\Gamma}^{*}_{\varkappa'\varrho''s}=0 \tag{18}$$

ist. Aus dem expliziten Ausdruck für $\Gamma^{*}_{\varkappa'\varrho''s}$ folgt aus (18)

$$\frac{1}{2}\left(\frac{\partial \bar{g}_{\varkappa'\varrho''}}{\partial \bar{x}^s}+\frac{\partial \bar{g}_{\varrho''s}}{\partial \bar{x}^{\varkappa'}}-\frac{\partial \bar{g}_{\varkappa's}}{\partial \bar{x}^{\varrho''}}\right)-\bar{C}_{\varkappa'\varrho''m}\,\bar{G}^m_s-C_{\varrho''sm}\,\bar{G}^m_{\varkappa'}+\bar{C}_{\varkappa'sm}\,\bar{G}^m_{\varrho''}=0. \tag{19}$$

Setzen wir in (19)

$$s = \mu'$$

und beachten (14), (17) und (17'), so kommt

$$\frac{\partial \bar{g}_{\varkappa'\mu'}}{\partial \bar{x}^{\varrho''}} = 0. \tag{20}$$

Auf entsprechende Weise folgt

$$\frac{\partial \bar{g}_{\varkappa''\mu''}}{\partial \bar{x}^{\varrho'}} = 0. \tag{20'}$$

Aus Gleichung (a) folgt endlich

$$\bar{L}^2(\bar{x}^1, \ldots, \bar{x}^n, \bar{v}^1, \ldots, \bar{v}^n) = \bar{L}_1^2(\bar{x}^1, \ldots, \bar{x}^r, \bar{v}^1, \ldots, \bar{v}^r) + \\ + \bar{L}_2^2(\bar{x}^{r+1}, \ldots, \bar{x}^n, \bar{v}^{r+1}, \ldots, \bar{v}^n).$$

Wir haben damit bewiesen

Satz 1. *Die notwendigen und hinreichenden Bedingungen dafür, daß ein Finslerscher Raum F_n als Produktraum eines F_r und F_{n-r} darstellbar ist, bestehen in der Existenz eines kovarianten positiv semidefiniten symmetrischen Vektorfeldes p'_{ik} vom Range r, für das*

$$Dp'_{ik} = 0 \tag{1}$$

ist, und die Kommutatorbildung sowohl bezüglich des durch p'_{ik} bestimmten r-dimensionalen Vektorraumes als auch bezüglich seines Normalraumes abgeschlossen ist.

Kehren wir jetzt zu dem Fall zurück, für den bloß das invariante Differential von p'_{ik} verschwindet.

Aus

$$Dp'_{ik} = 0 \tag{1}$$

folgt folgendes. Verschiebt man das orthonormierte r-Bein, das den Normalraum V_r von p'_{ik} aufspannt, längs zweier Linienelementfolgen in ein vorgegebenes Linienelement $(\underset{0}{x^i}, \underset{0}{v^i})$ parallel, so erhält man längs beiden Folgen zwei Tensoren q'_{ik} und r'_{ik}. Für beide gilt bei Differentiation längs den entsprechenden Folgen

$$\frac{D}{dt} q'_{ik} = \frac{D}{d\tau} r'_{ik} = 0.$$

Da aber (1) für ein gegebenes Linienelement ein eindeutiges Feld bestimmt, müssen die beiden Tensoren q'_{ik} und r'_{ik}, die ja im selben Linienelement definiert sind, zusammenfallen.

Aus

$$q'_{ik}(\underset{0}{x}, \underset{0}{v}) = r'_{ik}(\underset{0}{x}, \underset{0}{v}) = p'_{ik}(\underset{0}{x}, \underset{0}{v}) = \underset{0}{p'_{ik}}$$

folgt aber, daß die beiden orthonormierten r-Beine, die man durch die Parallelverschiebung erhält, in dem Vektorraum $\underset{0}{V_r}$ von $\underset{0}{p'_{ir}}$ liegen, und durch eine orthogonale Transformation verknüpft sind.

Demnach ist (1) hinreichend für die Existenz eines absolut parallelverschiebbaren Vektorraumes V_r. Setzen wir nun die Existenz eines absolut parallelverschiebbaren Vektorraumes voraus. Betrachten wir das Linienelement (x, v) und ein Weiteres $(\underset{0}{x}, \underset{0}{v})$. Verschiebt man $\underset{(\varrho')}{\xi}(x, r)$ längs zweier beliebiger Linienelementfolgen nach $(\underset{0}{x}, \underset{0}{v})$, so spannen sie stets die gleichen $\underset{0}{V_r}$ auf und das absolute Differential des von ihnen bestimmten symmetrischen Tensors muß immer Null sein. Es gilt also, wie im Falle der Riemannschen Geometrie:

Satz 2. *Die Existenz eines positiv semidefiniten symmetrischen Tensorfeldes mit verschwindendem invariantem Differential ist notwendig und hinreichend dafür, daß ein absolut parallelverschiebbarer r-dimensionaler Vektorraum V_r existiert.*

Aus Satz 2 folgt nun ebenso wie in der Riemannschen Geometrie, daß in diesem Falle die zu einem Linienelement gehörende Holonomiegruppe des Finslerschen Raumes reduzibel ist, falls die Länge eines Kurvenbogens nur von dem Bogen selbst, hingegen nicht von seiner Orientierung abhängt. Dies trifft dann zu, falls die Grundfunktion die starke Monodromiebedingung

$$L(x, -v) = L(x, v)$$

erfüllt. Aus dieser Bedingung folgt unmittelbar, daß auch die Parallelübertragung von Vektoren unabhängig von der Orientierung desjenigen Bogens ist, längs dem das, der Parallelverschiebung zugrunde liegende Linienelementfeld erklärt ist. Auf Grund dieser Bemerkung erledigt sich aber unsere Behauptung wortwörtlich wie im Falle der Riemannschen Geometrie.

(Eingegangen am 17. Dezember 1959.)

СКАЛЯРНЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ ИНВАРИАНТЫ ТРЕТЬЕГО ПОРЯДКА РИМАНОВОГО ПРОСТРАНСТВА ТРЕХ ИЗМЕРЕНИЙ

П. И. ПЕТРОВ (Казань, СССР)

(Представлено О. Варга)

1. Как известно, римановым пространством V_n называется дифференцируемое многообразие, метризованное с помощью неособенной дифференциальной квадратичной формы

$$ds^2 = g_{ij} dx^i dx^j. \tag{1}$$

Функции g_{ij} предполагаются дифференцируемыми нужное число раз. Дифференциальный инвариант риманового пространства — инвариант его основной метрической формы. В частности, под абсолютным скалярным дифференциальным инвариантом третьего порядка риманового многообразия разумеется скалярная функция

$$I\left(g_{\alpha\beta}, \frac{\partial g_{\alpha\beta}}{\partial x_i}, \frac{\partial^2 g_{\alpha\beta}}{\partial x_i \partial x_j}, \frac{\partial^3 g_{\alpha\beta}}{\partial x_i \partial x_j \partial x_k}\right), \tag{2}$$

неизменяющаяся при любых обратимых преобразованиях координат.

Цель настоящей статьи — построить базис полной системы [1] абсолютных скалярных дифференциальных инвариантов третьего порядка риманового пространства V_3.

2. С помощью теоремы приведения (см. [1], стр. 103—110) задача разыскания инвариантов (2) сводится к отысканию совокупных алгебраических инвариантов метрического тензора g_{ij}, его тензора кривизны R_{hijk} и ковариантной производной $R_{hijk,l}$ последнего. Так как в трехмерном пространстве тензор кривизны R_{hijk} выражается известной формулой

$$R_{\alpha\beta\gamma\delta} = g_{\beta\delta} L_{\alpha\gamma} - g_{\beta\gamma} L_{\alpha\delta} + g_{\alpha\gamma} L_{\beta\delta} - g_{\alpha\delta} L_{\beta\gamma}, \tag{3}$$

где L_{ik} — симметрический тензор, то последняя задача, в свою очередь, приводится к нахождению в пространстве трех измерений совокупных инвариантов следующих тензоров:

$$g_{\alpha\beta}, \quad L_{\alpha\beta}, \quad L_{\alpha\beta,\gamma}. \tag{4}$$

Компоненты тензора $L_{\alpha\beta,\gamma}$ удовлетворяют соотношениям

$$L^{\tau}{}_{\tau,\alpha} - L_{\tau\alpha,}{}^{\tau} = 0. \tag{5}$$

3. Применение установленной И. А. Схоутеном формулы инвариантного разложения аффинора валентности три на сумму упорядоченных элементарных величин дает

$$L_{\alpha\beta,\gamma} = L_{(\alpha\beta,\gamma)} + \frac{2}{3}\left(L_{(\alpha\beta),\gamma} - L_{\gamma(\alpha,\beta)}\right). \tag{6}$$

Первый из неприводимых относительно аффинной группы тензоров, фигурирующих в правой части равенства (6)

$$C_{\alpha\beta\gamma} = L_{(\alpha\beta,\gamma)} \tag{7}$$

является симметрическим по всем своим индексам и поэтому в трехмерном пространстве имеет десять существенных определяющих. Второй из них

$$T_{\alpha\beta\gamma} = \frac{2}{3}\left(L_{(\alpha\beta),\gamma} - L_{\gamma(\alpha,\beta)}\right) \tag{8}$$

удовлетворяет алгебраическим тождествам

$$T_{[\alpha\beta]\gamma} = 0, \quad T_{(\alpha\beta\gamma)} = 0, \tag{9}$$

вытекающим непосредственно из алгебраического свойства упорядоченных элементарных величин. Кроме того, соотношения (5) влекут за собою равенства

$$T_{\alpha\sigma}{}^{\sigma} = 0. \tag{10}$$

Следовательно, тензор $T_{\alpha\beta\gamma}$, определенный при помощи (8), имеет $27-(9+10+3)=5$ линейно независимых определяющих. Тензор $T_{\alpha\beta\gamma}$ можно отождествить с симметрическим тензором валентности два равным нулю следом. Для этого достаточно положить

$$T_{\alpha\beta} = T_{\alpha\lambda\mu}\,\varepsilon^{\lambda\mu}{}_{\beta}, \tag{11}$$

где $\varepsilon^{\lambda\mu\nu}$ — дискриминантный тензор пространства V_3.

Таким образом получается

Лемма 1. В пространстве трех измерений тензор $L_{\alpha\beta,\gamma}$ допускает два ортогональных линейных комитанта $C_{\alpha\beta\gamma}$, $T_{\alpha\beta}$, где $T^{\sigma}{}_{\sigma}=0$, являющихся, соответственно, симметрическими тензорами валентности три и два с 10 и 5 существенными определяющими.

В дальнейшем мы воспользуемся также следующей вспомогательной теоремой, доставляющей аналитический признак конформно-плоских V_3:

Лемма 2. V_3 тогда и только тогда конформно-плоское C_3, если $T_{\alpha\beta}=0$.

4. Базис полной системы абсолютных скалярных дифференциальных инвариантов третьего порядка трехмерного пространства Римана, который

по Т. И. Томасу обозначается символом $\mathfrak{S}(3, 3)$, состоит из 18 инвариантов. Базисные инварианты можно считать рациональными. Каждый абсолютный рациональный инвариант является частным двух целых рациональных инвариантов одинакового веса. Следовательно, искомый базис $\mathfrak{S}(3, 3)$ может быть образован при помощи 19 независимых относительных инвариантов нижеследующей серии симметрических тензоров

$$g_{\alpha\beta}, \quad L_{\alpha\beta}, \quad T_{\alpha\beta}, \quad C_{\alpha\beta\gamma}. \tag{12}$$

a) Инварианты тензора L_{ik}:

$$\psi_0 = 3!\, g_{[1\,[1}g_{22}g_{3]\,3]},\ [1]$$

$$\psi_1 = 3!\,(g_{[1\,[1}g_{22}L_{3]\,3]} + g_{[1\,[1}L_{22}g_{3]\,3]} + L_{[1\,[1}g_{22}g_{3]\,3]}),$$

$$\psi_2 = 3!\,(g_{[1\,[1}L_{22}L_{3]\,3]} + L_{[1\,[1}g_{22}L_{3]\,3]} + L_{[1\,[1}L_{22}g_{3]\,3]}),$$

$$\psi_3 = 3!\, L_{[1\,[1}L_{22}L_{3]\,3]}.$$

b) Инварианты тензора T_{ik}:

$$\varphi_1 = 3!\,(g_{[1\,[1}T_{22}T_{3]\,3]} + T_{[1\,[1}g_{22}T_{3]\,3]} + T_{[1\,[1}T_{22}g_{3]\,3]}),$$

$$\varphi_2 = 3!\, T_{[1\,[1}T_{22}T_{3]\,3]}.$$

c) Инварианты тензора $C_{\alpha\beta\gamma}$. Тернарная кубическая форма, отвечающая этому тензору, имеет два проективных инварианта

$$S = (abc)(abd)(acd)(bcd),$$

$$T = (abc)(abd)(ace)(bcf)(def)^2,$$

где $a, b, c, \ldots$ обозначают эквивалентные ряды символов.

Для образования метрических инвариантов тензора $C_{\alpha\beta\gamma}$ будем исходить из двух указанных ниже ортогональных комитантов последнего

$$a_{\alpha\beta} = C_{\alpha\sigma\tau}C^{\sigma\tau}{}_{\beta}, \quad b_{\alpha\beta} = C_{\alpha\beta\sigma}C^{\sigma\tau}{}_{\tau}.$$

Из способа образования тензоров $a_{\alpha\beta}, b_{\alpha\beta}$ ясно, что они симметрические. Рассматривая пару тернарных квадратичных форм, отвечающую тензорам $a_{\alpha\beta}, g_{\alpha\beta}$ можно наметить сразу два инварианта, соответственно, второй и четвертой степени относительно определяющих $C_{\alpha\beta\gamma}$:

$$I_3 = 3!\,(a_{[1\,[1}g_{22}g_{3]\,3]} + g_{[1\,[1}a_{22}g_{3]\,3]} + g_{[1\,[1}g_{22}a_{3]\,3]}),$$

$$I_4 = 3!\,(a_{[1\,[1}a_{22}g_{3]\,3]} + a_{[1\,[1}g_{22}a_{3]\,3]} + g_{[1\,[1}a_{22}a_{3]\,3]}).$$

[1] Под выражением $a_{[i_1\,[k_1}a_{i_2k_2}\cdots a_{i_m]\,k_m]}$ понимают, как обычно, тензор, образованный двукратным применением операции альтернирования к произведению $a_{i_1k_1}a_{i_2k_2}\cdots$ $\cdots a_{i_mk_m}$, причем альтернируется ряд индексов $i_1, i_2, \ldots, i_m$ и независимо от него ряд индексов $k_1, k_2, \ldots, k_m$.

Подобным же образом находим еще два инварианта:

$$I_5 = 3!\,(b_{[1\,[1}g_{22}g_{3]\,3]} + g_{[1\,[1}b_{22}g_{3]\,3]} + g_{[1\,[1}g_{22}b_{3]\,3]}),$$

$$I_6 = 3!\,(b_{[1\,[1}b_{22}g_{3]\,3]} + b_{[1\,[1}g_{22}b_{3]\,3]} + g_{[1\,[1}b_{22}b_{3]\,3]}).$$

Наконец, возьмем совокупный инвариант трех тензоров

$$I_7 = 3!\,(a_{[1\,[1}b_{22}g_{3]\,3]} + b_{[1\,[1}a_{22}g_{3]\,3]} + a_{[1\,[1}g_{22}b_{3]\,3]} + b_{[1\,[1}g_{22}a_{3]\,3]} + g_{[1\,[1}a_{22}b_{3]\,3]} + g_{[1\,[1}b_{22}a_{3]\,3]}).$$

d) Совокупные инварианты тензоров T_{ik}, $C_{\alpha\beta\gamma}$:

$$\omega_1 = 3!\,g_{[1\,[1}a_{22}T_{3]\,3]} + g_{[1\,[1}T_{22}a_{3]\,3]} + a_{[1\,[1}g_{22}T_{3]\,3]} + T_{[1\,[1}g_{22}a_{3]\,3]} + a_{[1\,[1}T_{22}g_{3]\,3]} + T_{[1\,[1}a_{22}g_{3]\,3]}),$$

$$\omega_2 = 3!\,(g_{[1\,[1}b_{22}T_{3]\,3]} + g_{[1\,[1}T_{22}b_{3]\,3]} + b_{[1\,[1}g_{22}T_{3]\,3]} + T_{[1\,[1}g_{22}b_{3]\,3]} + b_{[1\,[1}T_{22}g_{3]\,3]} + T_{[1\,[1}b_{22}g_{3]\,3]}),$$

$$\omega_3 = T_{\alpha\beta}u^\alpha u^\beta, \quad u_i = C_{i\sigma}{}^{\sigma}.$$

e) Совокупные инварианты тензоров L_{ik}, $C_{\alpha\beta\gamma}$:

$$\bar{\omega}_1 = 3!\,(g_{[1\,[1}a_{22}L_{3]\,3]} + g_{[1\,[1}L_{22}a_{3]\,3]} + a_{[1\,[1}g_{22}L_{3]\,3]} + L_{[1\,[1}g_{22}a_{3]\,3]} + a_{[1\,[1}L_{22}g_{3]\,3]} + L_{[1\,[1}a_{22}g_{3]\,3]}),$$

$$\bar{\omega}_2 = 3!\,(g_{[1\,[1}b_{22}L_{3]\,3]} + g_{[1\,[1}L_{22}b_{3]\,3]} + b_{[1\,[1}g_{22}L_{3]\,3]} + L_{[1\,[1}g_{22}b_{3]\,3]} + b_{[1\,[1}L_{22}g_{3]\,3]} + L_{[1\,[1}b_{22}g_{3]\,3]}),$$

$$\bar{\omega}_3 = L_{\alpha\beta}u^\alpha u^\beta.$$

5. Остается еще проверить функциональную независимость построенных выше целых рациональных инвариантов относительно переменных $C_{\alpha\beta\gamma}$, T_{ij}, L_{ik}. Чтобы решить этот на вид простой вопрос нужно выполнить ряд технически сложных выкладок. Упростить встречающиеся при этом вычисления помогут нам два простых замечания, составляющие основу применяемого здесь нами метода. Значение его не исчерпывается приложением к интересующей нас задаче, так как он может быть успешно использован для решения широкого круга подобных же, но более трудных проблем. Поскольку количество независимых функций в заданной совокупности определяется рангом якобиевой матрицы этих функций, то и наши замечания будут касаться вопроса об обнаружении максимальности ее ранга. С этой целью нам достаточно убедиться в максимальности ранга для какой-нибудь подходящим образом выбранной системы числовых значений независимых переменных (эту совокупность числовых значений ком-

понент тензоров C_{ijk}, T_{ik}, L_{ij} условимся впредь называть коротко точкой M^0). Действительно, любой минор $\Delta_{i/M^0} \neq 0$ якобиевой матрицы совокупности полиномов от многих переменных, будучи непрерывной функцией, сохранит свой знак и во всех точках некоторой окрестности $U(M^0)$ точки M^0. А это будет означать функциональную независимость рассматриваемых функций в области $U(M^0)$. Если требуется вычислить производную

$$\frac{\partial F(t^1, t^2, \ldots, t^n)}{\partial t^k}$$

при аргументах, обращающихся в постоянные значения t_0^i, то при этом вычислении все аргументы кроме одного t^k можно заранее положить равными $t_0^1, t_0^2, \ldots, t_0^{k-1}, t_0^{k+1}, \ldots, t_0^n$ и уже затем находить производную. Из только что высказанного второго замечания непосредственно вытекают: моном, у которого два множителя в точке M^0 обращаются в нуль, не оказывает никакого влияния на значения частных производных инварианта в этой точке; моном, входящий как слагаемое в состав инварианта и содержащий один единственный обращающийся в опорной точке M^0, в нуль множитель, можно в целях вычисления частной производной от этого инварианта заменить на линейную функцию относительно этого множителя — переменного.

Руководствуясь сформулированными замечаниями, нетрудно доказать независимость целых рациональных функций

$$\psi_1, \psi_2, \psi_3, \varphi_1, \varphi_2, I_1 = \frac{3}{2} S, I_2 = 4T, I_3, \ldots, I_7, \omega_1, \omega_2, \omega_3, \overline{\omega}_1, \overline{\omega}_2, \overline{\omega}_3.$$

При этом мы полагаем

$$g_{ii} = \varepsilon_i = \pm 1, \quad g_{ik} = 0 \qquad (i \neq k)$$

и воспользуемся обозначениями

$$\begin{aligned} &C_{111} = x_1, \quad C_{122} = x_4, \quad C_{233} = x_7, \quad C_{333} = x_{10}; \\ &C_{112} = x_2, \quad C_{133} = x_5, \quad C_{222} = x_8, \\ &C_{113} = x_3, \quad C_{123} = x_6, \quad C_{223} = x_9, \end{aligned}$$

$$\begin{aligned} &T_{11} = y_1, \quad T_{22} = y_2, \quad T_{33} = y_3, \\ &T_{12} = y_4, \quad T_{13} = y_5, \quad T_{23} = y_6; \end{aligned}$$

$$\begin{aligned} &L_{11} = z_1, \quad L_{22} = z_2, \quad L_{33} = z_3, \\ &L_{12} = z_4, \quad L_{13} = z_5, \quad L_{23} = z_6. \end{aligned}$$

Якобиев определитель

$$I_{18} = \frac{\partial (I_1 I_2 \cdots I_7 \, \varphi_1 \varphi_2 \omega_1 \omega_2 \omega_3 \, \psi_1 \psi_2 \psi_3 \overline{\omega}_1 \overline{\omega}_2 \overline{\omega}_3)}{\partial (x_5 x_{10} \, x_1 x_2 x_3 x_7 x_9 \, y_1 y_2 y_4 y_5 y_6 \, z_1 z_2 z_3 z_4 z_5 z_6)}$$

для фиксированных значений переменных

$$x_1^0=1,\quad x_2^0=1,\quad x_3^0=x_4^0=x_5^0=x_6^0=x_7^0=x_8^0=0,\quad x_9^0=1,\quad x_{10}^0=0,$$

$$y_1^0=\varepsilon_1,\quad y_2^0=-\varepsilon_2,\quad y_4^0=y_5^0=y_6^0=0,$$

$$z_1^0=1,\quad z_2^0=2,\quad z_3^0=z_4^0=z_5^0=z_6^0=0$$

имеет вид

$$I_{18/M^0}=\begin{vmatrix} \Delta_7 & 0 & 0 \\ * & \Delta_5 & 0 \\ * & * & \Delta_6 \end{vmatrix},$$

где

$$\Delta_7=\begin{vmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ & & 2\varepsilon_2\varepsilon_3 & 6\varepsilon_1\varepsilon_3 & 0 & 0 & 6\varepsilon_1\varepsilon_2 \\ & * & 6\varepsilon_2 & 8\varepsilon_1\varepsilon_2\varepsilon_3+14\varepsilon_1 & 0 & 0 & 8\varepsilon_1\varepsilon_2\varepsilon_3+14\varepsilon_1+6\varepsilon_2 \\ & & 2\varepsilon_2\varepsilon_3 & 2\varepsilon_1\varepsilon_3 & 2 & 2 & 2\varepsilon_1\varepsilon_2 \\ & & 2\varepsilon_2-2\varepsilon_3 & -2\varepsilon_3 & \varepsilon_1+\varepsilon_2+\varepsilon_3 & \varepsilon_1+\varepsilon_2-2\varepsilon_3 & 2\varepsilon_2 \\ & & 8\varepsilon_2-2\varepsilon_3 & 4\varepsilon_1\varepsilon_2\varepsilon_3+10\varepsilon_1-2\varepsilon_3 & \varepsilon_1+\varepsilon_2+4\varepsilon_3 & \varepsilon_1+4\varepsilon_2+\varepsilon_3 & 10\varepsilon_1+8\varepsilon_2+4\varepsilon_1\varepsilon_2\varepsilon_3 \end{vmatrix},$$

$$\Delta_5=\begin{vmatrix} -\varepsilon_2\varepsilon_3 & \varepsilon_1\varepsilon_3 & 0 & 0 & 0 \\ \varepsilon_2\varepsilon_3 & \varepsilon_1\varepsilon_3 & 0 & 0 & 0 \\ & & -2\varepsilon_3 & 0 & 0 \\ & * & -2\varepsilon_3 & 0 & -2\varepsilon_2 \\ & & * & 2\varepsilon_2\varepsilon_3 & * \end{vmatrix},$$

$$\Delta_6=\begin{vmatrix} \varepsilon_2\varepsilon_3 & \varepsilon_1\varepsilon_3 & \varepsilon_1\varepsilon_2 & 0 & 0 & 0 \\ 2\varepsilon_3 & \varepsilon_3 & 2\varepsilon_1+\varepsilon_2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ & & & -2\varepsilon_3 & 0 & 0 \\ & * & & -2\varepsilon_3 & 0 & -2\varepsilon_2 \\ & & & * & 2\varepsilon_2\varepsilon_3 & * \end{vmatrix}.$$

В точке M^0 ни один из определителей Δ_7, Δ_5, Δ_6 не обращается в нуль. Вместе с тем $I_{18}\neq 0$ во всех точках некоторой окрестности точки $M^0(x_i^0, y_k^0, z_j^0)$, следовательно, инварианты в этой области независимы между собой.

Подводя итог всему исследованию, можем формулировать предложение [2]:

Теорема 1. Инварианты

$$g=|g_{ik}|,\quad I_i,\quad \varphi_k,\quad \psi_j,\quad \omega_j,\quad \overline{\omega}_j \qquad (i=1,\ldots,7;\quad j=1,2,3;\quad k=1,2)$$

образуют базис фундаментальной системы скалярных дифференциальных инвариантов третьего порядка риманового пространства трех измерений V_3 с произвольной сигнатурой.

Отметим следующие теоремы, вытекающие из теоремы 1 и представляющие сами по себе интерес:

Теорема 2. Базис функционально независимых скалярных дифференциальных инвариантов третьего порядка пространства C_3 состоит из 14 инвариантов

$$g,\quad \psi_j,\quad I_i,\quad \overline{\omega}_j.$$

Теорема 3. Скалярные функции

$$g,\quad I_i \qquad (i=1,\ldots,7)$$

образуют базис функционально независимых метрических инвариантов тернарной кубической формы.

Теорема 4. Какова бы ни была сигнатура неособенной тернарной дифференциальной квадратичной формы

$$ds^2=g_{ij}dx^i dx^j,$$

инварианты

$$g,\quad \psi_1,\quad \psi_2,\quad \psi_3$$

образуют базис полной системы ее скалярных дифференциальных инвариантов второго порядка.

Основные выводы работ [3], [4] об инвариантах содержатся как частные случаи в теореме 4.

Институт математики
Казанского университета

(Поступило 31. VIII. 1959.)

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] T. Y. Thomas, *The differential invariants of generalized spaces* (Cambridge, 1934), § 75.
[2] П. И. Петров, Дифференциальные инварианты пространств Римана, ДАН СССР, **58** (1947), стр. 1273—1274.
[3] E. B. Christoffel, Über die Transformation der homogenen Differentialausdrücke zweiten Grades, *Journal f. reine u. angew. Math.*, **70** (1869), стр. 67.
[4] Ф. М. Суворов, О характеристиках систем трех измерений (Казань, 1871).

ON TURÁN'S FIRST MAIN THEOREM

By

N. G. DE BRUIJN (Amsterdam)

(*Presented by* P. TURÁN)

1. Introduction. Recently, E. MAKAI[1] determined the best-possible constant in a theorem of TURÁN. The theorem is quoted as TURÁN's first main theorem, since it appeared in his book[2] as the solution of his "first problem". It will be cited in Sec. 4 below.

This best-possible constant was obtained independently (though somewhat later) by the present author, whose proof appeared to be considerably shorter that the one of Mr. MAKAI. The matter now seems to have been rounded off so nicely that it is worth while to present a self-contained account of it. The arrangement is as follows. In Sec. 3 we present a new proof of an elegant theorem, due to MAKAI.[3] Sec. 4 gives the proof of the main theorem, and Sec. 6 the new proof of the fact that the constant is best-possible. Finally, Sec. 7 gives a discussion of the cases where the equality sign occurs.

It should be remarked that Sec. 4 was implicitly contained in the paper of I. DANCS,[4] which inspired both Mr. MAKAI and the present author.

2. Notation. Throughout the paper m and k are integers, $m \geqq 0$, $k \geqq 1$. And $z_1, \ldots, z_k$ are complex numbers, with $|z_1| \geqq 1, \ldots, |z_k| \geqq 1$. We put $z_j = -(u_j)^{-1}$, whence $|u_j| \leqq 1$. And we write

$$f(z) = \prod_{j=1}^{k} (1 + z u_j) = \prod_{j=1}^{k} (1 - z/z_j),$$

$$(f(z))^{-1} = a_0 + a_1 z + a_2 z^2 + \ldots,$$

$$h_m(z) = (-1)^{m-1} \left\{ 1 - f(z) \sum_{\mu=0}^{m} a_\mu z^\mu \right\}.$$

1 E. MAKAI, The first main theorem of P. Turán, *Acta Math. Acad. Sci. Hung.*, 10 (1959), pp. 405—411.

2 P. TURÁN, *Eine neue Methode in der Analysis und deren Anwendungen* (Budapest, 1953).

3 E. MAKAI, On a maximum problem, *Acta Math. Acad. Sci. Hung.*, 9 (1958), pp. 105—110.

4 I. DANCS, On an extremal problem, *Acta Math. Acad. Sci. Hung.*, 9 (1958), pp. 309—313.

It is easy to see that $h_m(z)$ has the form

$$h_m(z)=\sum_{\nu=m+1}^{m+k}c_\nu z^\nu.$$

Though the notation does not reveal it, the functions f, h_m, and the coefficients a_μ, c_ν depend on $z_1,\dots,z_k$. In the special case where $u_1=\dots=u_k=1$ (so $z_1=\dots=z_k=-1$) we shall write $f_1^*, h_m^*, a_\mu^*, c_\nu^*$. So $f^*(z)=(1+z)^k$, $a_\mu^*=(-1)^\mu\binom{k+\mu-1}{\mu}$. In particular we shall consider $N_{m,k}$, defined by

$$N_{m,k}=h_m^*(1)=c_{m+1}^*+\dots+c_{m+k}^*.$$

3. Theorem of E. Makai. *For each ν, the coefficient c_ν can be expressed as a (symmetrical) polynomial in $u_1,\dots,u_k$, with non-negative coefficients.*

Proof. We use a generating function, introducing a variable t. We have

$$\sum_{m=0}^{\infty}h_m(z)t^m=\sum_{m=0}^{\infty}(-1)^{m-1}t^m+\sum_{m=0}^{\infty}(-1)^m t^m f(z)\sum_{\mu=0}^{m}a_\mu z^\mu=$$

$$=-\frac{1}{1+t}+f(z)\sum_{\mu=0}^{\infty}a_\mu(-tz)^\mu(1-t+t^2-\cdots)=$$

$$=\frac{-1}{1+t}+\frac{f(z)}{(1+t)f(-tz)}=\frac{1}{1+t}\left\{-1+\prod_{j=1}^{k}\left(1+\frac{(1+t)u_jz}{1-u_jzt}\right)\right\}.$$

It is now sufficient to remark (i) that

$$(1+t)^{-1}\left\{-1+\prod_{j=1}^{k}(1+(1+t)x_j)\right\}$$

is a polynomial in $t, x_1,\dots,x_k$, with non-negative coefficients, and (ii) that $x_j=u_jz(1-u_jzt)^{-1}$ can be expanded in terms of u_j, z, t with non-negative coefficients.

Corollary. We have $c_\nu^*\geqq 0$, and, by virtue of $|u_1|\leqq 1,\dots,|u_n|\leqq 1$, we have

(3. 1) $$|c_{m+1}|+\dots+|c_{m+k}|\leqq c_{m+1}^*+\dots+c_{m+k}^*=N_{m,k}.$$

4. Turán's first main theorem. We have (see Sec. 2)

$$1-(-1)^{m-1}\sum_{\nu=m+1}^{m+k}c_\nu z^k=f(z)\sum_{\mu=0}^{m}a_\mu z^\mu.$$

It follows that the left-hand side vanishes if z has one of the values $z_1,\dots,z_k$. On substitution of $z=z_j$ and multiplication by a complex variable b_j, we

obtain, after summation with respect to j, the identity

$$(4.1) \qquad b_1+\cdots+b_k=(-1)^{m-1}\sum_{\nu=m+1}^{m+k} c_\nu \sum_{j=1}^{k} b_j z_j^\nu .$$

By (3.1), we infer that $|z_1|\geqq 1,\ldots,|z_k|\geqq 1$ imply that

$$(4.2) \qquad |b_1+\cdots+b_k|\leqq N_{m,k} \max_{m+1\leqq\nu\leqq m+k}\left|\sum_{j=1}^{k} b_j z_j^\nu\right|.$$

5. Evaluation of $N_{m,k}$. We immediately have (cf. Sec. 2)

$$(5.1) \qquad N_{m,k}=h_m^*(1)=(-1)^{m-1}\left\{1-2^k\sum_{\mu=0}^{m}(-1)^\mu\binom{k+\mu-1}{\mu}\right\}=$$

$$=2^k\left\{\binom{k+m-1}{m}-\binom{k+m-2}{m-1}+\cdots+(-1)^m\binom{k-1}{0}\right\}+(-1)^{m-1}.$$

Mr. MAKAI obtained a different expression for the best-possible constant (see [1]), which we derive here from the generating function (cf. Sec. 3):

$$\sum_{m=0}^{\infty} N_{m,k}t^m=\sum_{m=0}^{\infty} h_m^*(1)t^m=(1+t)^{-1}\left\{-1+\left(1+\frac{1+t}{1-t}\right)^k\right\}=$$

$$=(1-t)^{-k}\,\frac{2^k-(1-t)^k}{2-(1-t)}=\sum_{j=0}^{k-1}2^j(1-t)^{-j-1},$$

whence

$$(5.2) \qquad N_{m,k}=\sum_{j=0}^{k-1}2^j\binom{j+m}{m}.$$

E. MAKAI (see [1]) shows that this expression is equal to $N_{m,k}=$ $=\binom{m+k}{k}k\int_0^1 x^m(1+x)^{k-1}\,dx$, but this formula plays no role in the present paper.

6. Proof that $N_{m,k}$ is best-possible. Let m and k be given. We shall show that, in (4.2), the number $N_{m,k}$ cannot be replaced by any smaller number which is still independent of $b_1,\ldots,b_k,z_1,\ldots,z_k$. In other words, we shall show that for every number p with $0<p<N_{m,k}$, there exist complex numbers $b_1,\ldots,b_k,z_1,\ldots,z_k$ with

$$(6.1) \qquad |b_1+\cdots+b_k|>p \max_{m+1\leqq\nu\leqq m+k}\left|\sum_{j=1}^{n} b_j z_j^\nu\right|.$$

PROOF. We have $h_m^*(1)=N_{m,k}$, and $h_m^*(1)$ is the special value of $h_m(1)=\sum_{\nu=m+1}^{m+k} c_\nu$ in the case that $z_1=\cdots=z_k=-1$. We notice that $h_m(1)$

depends continuously on $z_1, \ldots, z_k$. It follows that there is a positive number δ such that $|h_m(1)| > p$ whenever $|z_1+1| < \delta, \ldots, |z_k+1| < \delta$. We now fix $z_1, \ldots, z_k$ according to these inequalities, with the only further condition that $z_1, \ldots, z_k$ are mutually different.

We next take $b_1, \ldots, b_k$ such that

$$\sum_{j=1}^{k} b_j z_j^{\nu} = 1 \qquad (\nu = m+1, \ldots, m+k).$$

This set of k linear equations has a unique solution for $b_1, \ldots, b_k$. (Its determinant is a Vandermonde determinant and does not vanish, the z_j's being mutually different.)

By (4. 1) we now have

$$|b_1 + \cdots + b_k| = \left| (-1)^{m-1} \sum_{\nu=m+1}^{m+k} c_\nu \right| = |h_m(1)| > p,$$

and (6. 1) follows.

7. Discussion of the equality sign. We first discuss the equality sign in (3. 1). We show that we have[5] $|c_{m+1}| = c^*_{m+1}$ if and only if $u_1 = \cdots = u_k$, $|u_1| = 1$. If $k=1$, this is trivial, c_{m+1} being a constant (positive) multiple of a power of u_1 in that case. If $k>1$, we can remark that the polynomial c_{m+1} contains a term $a u_1^{m+1}$ as well as a term $b u_1^m u_2$, with $a>0$, $b>0$. As $|u_1| \leqq 1$, $|u_2| \leqq 1$, we infer from $|a u_1^{m+1} + b u_1^m u_2| = a+b$ that $u_1 = u_2$, $|u_1| = 1$. Replacing the index 2 by any other index, our statement follows.

Conversely, if $u_1 = \cdots = u_k$, $|u_1| = 1$, it is easy to see that we have equality in (3. 1).

Next we consider (4. 2). If $k=1$, the equality sign holds as soon as $|z_1| = 1$. So we can fix our attention to $k>1$. Moreover, we assume that $b_1 + \cdots + b_k \neq 0$. Denoting the max on the right-hand side of (4. 2) by M, we notice that $M>0$, the left-hand side being positive. Assuming equality in (4. 2), and observing how (4. 2) followed from (4. 1), we deduce that $|c_{m+1} M| = c^*_{m+1} M$, whence $|c_{m+1}| = c^*_{m+1}$. It follows that $u_1 = \cdots = u_k$, $|u_1| = 1$.

But in this case we have no equality in (4. 2), as can be seen, for example, from the fact that $N_{m,k} \neq 1$ ((5. 1) gives $N_{m,k} \geqq 2(m+1) > 1$, using $j=1$ only).

Our conclusion is that if $k>1$, the equality sign in (4. 2) holds only if both sides vanish.

(Received 13 November 1959)

[5] In Dancs' paper (see [4], p. 310) it seems to be asserted that the statement made here for c_{m+1}, equally holds for $c_{m+1}, \ldots, c_{m+k}$. It is not proved there, however, and anyway it fails for c_{m+k}.

HALBGRUPPEN UND RINGE MIT LINKSEINHEITEN OHNE LINKSEINSELEMENTE

Von

L. RÉDEI (Szeged), Mitglied der Akademie

Herr P. TURÁN hat mir freundlich das Problem aufgeworfen, ob es Ringe mit Linkseinheiten ohne Linkseinselemente gibt, wobei man unter einer Linkseinheit einer Struktur S jedes Element $\varepsilon(\in S)$ mit $\varepsilon S = S$ zu verstehen hat, vorausgesetzt, daß in S eine Multiplikation definiert ist. Die Antwort ist bejahend, sogar hat man das Gefühl, daß sehr viele solche Ringe existieren, jedoch gelang uns nur mit Mühe, ein Beispiel dafür anzugeben.

Für reelle Zahlen x setzen wir

$$x^+ = \max(0, x).$$

Für eine Aussage A bezeichnet $[A]$ die Zahl 1 oder 0, je nachdem A richtig oder falsch ist. Unter einer Halbgruppe verstehen wir eine multiplikative assoziative Struktur.

SATZ. *Die durch die Gleichungen*

$$\varrho = \varrho^2\sigma, \quad \sigma = \varrho\sigma^2 \tag{1}$$

definierte Halbgruppe H hat ϱ zur Linkseinheit, hat aber kein Linkseinselement. Ähnliches gilt für den Halbgruppenring $R(H)$ von H über einem Ring R mit Einselement. Die Elemente von H lassen sich eindeutig in der Form

$$\sigma^a(\varrho\sigma)^b\varrho^c \qquad (a, b, c = 0, 1, \ldots; a+b+c > 0) \tag{2}$$

schreiben.[1] *Wird das Element* (2) *kurz mit* (a, b, c) *bezeichnet, so lautet die Multiplikation in H als*

$$(a, b, c)(r, s, t) = (a + (r-c)^+, b[r \leqq c] + s[r \geqq c] + [r = c > 0], (c-r)^+ + t). \tag{3}$$

BEMERKUNG. Die erste Hälfte des Satzes gilt auch für alle homomorphen Bilder H' von H ohne Linkseinselement, d. h. für jedes solche H' besitzen H' und $R(H')$ Linkseinheiten, aber keine Linkseinselemente. Solche homomorphen Bilder H' von H lassen sich mit Leichtigkeit angeben. Mit Erlaubnis von Herrn L. FUCHS betrachten wir das von ihm entdeckte besonders einfache Beispiel eines solchen H', das nämlich so entsteht, daß man den zwei Glei-

[1] Die Faktoren mit dem Exponenten 0 sind außer acht zu lassen.

chungen (1) die weitere definierende Gleichung $\sigma = \sigma\varrho\sigma$ hinzunimmt, d. h. man für H' die durch die Gleichungen

$$\varrho = \varrho^2\sigma, \quad \sigma = \varrho\sigma^2, \quad \sigma = \sigma\varrho\sigma$$

definierte Halbgruppe nimmt. Dann hat nämlich H' (und genau so auch $R(H')$) die Linkseinheit ϱ, hat sogar das Rechtseinselement $\varrho\sigma$, hat aber kein Linkseinselement. Es genügt das letztere zu beweisen. Das kann nach Herrn FUCHS sehr leicht so geschehen, daß man H' durch eine passende Halbgruppe von Abbildungen (treu oder untreu) repräsentiert. Und zwar betrachte man die Menge M der abzählbar unendlichen Folgen $(n_1, n_2, \ldots)$ von nichtnegativen ganzen Zahlen und bezeichne mit ϱ und σ die beiden Abbildungen

$$(n_1, n_2, \ldots) \to (0, n_1, n_2, \ldots), \quad (n_1, n_2, \ldots) \to (0, n_3, n_4, \ldots)$$

von M in sich. Man weist leicht nach, daß dann die obigen drei Gleichungen erfüllt sind, indem man das Produkt $\alpha\beta$ zweier Abbildungen α, β so deutet, daß man zuerst α, dann β ausführt. Es genügt zu beweisen, daß kein Element ξ der durch ϱ und σ erzeugten Halbgruppe ein Linkseinselement ist. Das ist aber klar, da das durch $\xi\varrho$ hervorgerufene Bild eines jeden Elementes von M von der Form $(0, 0, c, \ldots)$ ist, woraus $\xi\varrho \neq \varrho$ folgt. Ein ähnlich leichter Beweis des obigen allgemeineren Satzes ist nicht zu hoffen.

Zuerst beweisen wir, daß (2) die eindeutige Darstellung der Elemente von H ist und (3) besteht. Den Beweis führen wir so aus, daß wir mit H' die Menge aller Tripel

$$(a, b, c) \qquad (a, b, c = 0, 1, \ldots; a + b + c > 0) \tag{4}$$

bezeichnen, zeigen dann, daß in H' durch (3) eine assoziative Multiplikation definiert wird (d. h. H' eine Halbgruppe ist), ferner zeigen wir, daß H und H' isomorph sind.

Zunächst sehen wir leicht, daß auf der rechten Seite von (3) nicht alle drei „Komponenten" verschwinden, weshalb in H' durch (3) tatsächlich eine Multiplikation definiert ist.

Es ergeben sich die Spezialfälle von (3):[2]

$$\begin{aligned} (a, 0, 0)(r, 0, 0) &= (a + r, 0, 0) && (a + r > 0), \\ (0, b, 0)(0, s, 0) &= (0, b + s, 0) && (b + s > 0), \\ (0, 0, c)(0, 0, t) &= (0, 0, c + t) && (c + t > 0). \end{aligned}$$

Ferner folgt aus (3):

$$(a, b, c) = (a, b, 0)(0, 0, c) = ((a, 0, 0)(0, b, 0))(0, 0, c) \qquad (a + b + c > 0).$$

[2] Faktoren von der Form $(0, 0, 0)$ sind außer acht zu lassen.

Nach diesen Formeln gilt

$$(a, b, c) = ((1, 0, 0)^a (0, 1, 0)^b)(0, 0, 1)^c \qquad (a+b+c>0).$$

Somit wird H' durch $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ erzeugt. Da aber nach (3)

$$(0, 0, 1)(1, 0, 0) = (0, 1, 0)$$

ist, so wird H' sogar schon durch $(1, 0, 0), (0, 0, 1)$ erzeugt. Um also die Assoziativität von (3) auszuweisen, bekanntlich[3] genügt es

$$((a, b, c)(r, s, t))(1, 0, 0) = (a, b, c)((r, s, t)(1, 0, 0)), \tag{5}$$

$$((a, b, c)(r, s, t))(0, 0, 1) = (a, b, c)((r, s, t)(0, 0, 1)) \tag{6}$$

zu zeigen.

Zuerst beweisen wir (5). Nach (3) gilt

$$(a, b, c)(1, 0, 0) = (a+(1-c)^+, b[c>0]+[c=1], (c-1)^+), \tag{7}$$

also den Fällen $c=0$, $c=1$, $c>1$ entsprechend:

$$(a, b, 0)(1, 0, 0) = (a+1, 0, 0), \tag{8}$$

$$(a, b, 1)(1, 0, 0) = (a, b+1, 0), \tag{8'}$$

$$(a, b, c)(1, 0, 0) = (a, b, c-1) \qquad (c>1). \tag{8''}$$

Von hier an unterscheiden wir im Beweis von (5) drei Fälle:

Fall $t=0$. Die linke und rechte Seite von (5) ist jetzt nach (3) bzw. (8) gleich

$$(a+(r-c)^+, b[r\leqq c]+s[r\geqq c]+[r=c>0], (c-r)^+)(1, 0, 0)$$

bzw.

$$(a, b, c)(r+1, 0, 0).$$

Diese gehen nach (7) bzw. (3) in

$$(a+(r-c)^+ + (1-(c-r)^+)^+, (b[r\leqq c]+s[r\geqq c]+[r=c>0])[c>r]+ \\ +[c=r+1], ((c-r)^+-1)^+)$$

bzw.

$$(a+(r+1-c)^+, b[r+1\leqq c]+[r+1=c], (c-r-1)^+)$$

über. Je nach den Fällen $r=c$, $r<c$, $r>c$ sind beide Elemente gleich

$(a+1, 0, 0)$ bzw. $(a, b+[c=r+1], c-r-1)$ bzw. $(a+r-c+1, 0, 0)$.

Somit ist (5) in diesem Fall richtig.

[3] Vgl. L. Rédei, *Algebra I* (Leipzig, 1959), Satz 30.

Fall $t=1$. Die linke und rechte Seite von (5) ist jetzt nach (3) bzw. (8′) gleich

$$(a+(r-c)^+,\ b[r\leqq c]+s[r\geqq c]+[r=c>0],\ (c-r)^++1)\,(1,0,0)$$

bzw.

$$(a,b,c)\,(r,\ s+1,0).$$

Diese gehen nach (7) bzw. (3) in

$$(a+(r-c)^+,\ b[r\leqq c]+s[r\geqq c]+[r=c>0]+[r\geqq c],\ (c-r)^+)$$

bzw.

$$(a+(r-c)^+,\ b[r\leqq c]+(s+1)\,[r\geqq c]+[r=c>0],\ (c-r)^+)$$

über. Da diese zwei Elemente gleich sind, ist (5) im vorliegenden Fall richtig.

Fall $t\geqq 2$. Die linke Seite von (5) ist jetzt nach (8″) gleich

$$(a+(r-c)^+,\ b[r\leqq c]+s[r\geqq c]+[r=c>0],\ (c-r)^++t-1).$$

Die rechte Seite von (5) ist ebenfalls nach (8″) gleich

$$(a,b,c)\,(r,s,t-1).$$

Dies ist aber nach (3) dem vorigen Elemente gleich. Somit ist (5) bewiesen.

Der Beweis von (6) ist leicht. Da nämlich die rechte Seite von (3) von der Form $(u,v,w+t)$ mit von t unabhängigen u,v,w ist, da ferner nach (3)

$$(a,b,c)\,(0,0,1)=(a,b,c+1)$$

ist, so folgt hieraus (6) offenbar.

Bisher haben wir bewiesen, daß (die Multiplikation (3) assoziativ, d.h.) H' eine Halbgruppe ist. Wir wollen jetzt beweisen, daß H und H' isomorph sind.

Es wurde schon bemerkt, daß H' durch $(1,0,0)$ und $(0,0,1)$ erzeugt wird. Anderseits berechnen wir

$$(0,0,1)^2\,(1,0,0)=(0,0,2)\,(1,0,0)=(0,0,1),$$

$$(0,0,1)\,(1,0,0)^2=(0,0,1)\,(2,0,0)=(1,0,0).$$

Hiernach sind die Gleichungen (1) mit $(0,0,1)$ und $(1,0,0)$ an Stelle von ϱ und σ erfüllt. Das alles bedeutet, daß H durch

$$\varrho\to(0,0,1),\quad \sigma\to(1,0,0) \tag{9}$$

homomorph auf H' abgebildet wird. Wir zeigen gerade, daß dieser Homomorphismus ein Isomorphismus ist. Es genügt zu zeigen, daß bei unserem Homomorphismus verschiedene Elemente von H stets auf solche von H' abgebildet werden.

Zu diesem Zweck betrachten wir zunächst ein beliebiges Element ω von H. Dieses ist dann von der Form

$$\omega = \tau_1 \ldots \tau_n \qquad (n \geqq 1),$$

wobei jeder Faktor τ_i gleich ϱ oder σ ist. Unter allen solchen Darstellungen von ω nehmen wir eine mit minimaler Faktorenzahl n an. Anderseits schreiben wir (1) als

$$\varrho\varrho\sigma = \varrho, \quad \varrho\sigma\sigma = \sigma.$$

Hieraus und aus der Minimalität von n folgt, daß

$$\tau_i = \varrho, \quad \tau_{i+2} = \sigma \qquad (1 \leqq i \leqq n-2)$$

für kein i möglich ist, also im Fall $\tau_k = \tau_{k+1} = \varrho$ notwendig $\tau_k = \cdots = \tau_n = \varrho$ und im Fall $\tau_k = \tau_{k+1} = \sigma$ notwendig $\tau_1 = \cdots = \tau_{k+1} = \sigma$ sein muß. Also muß jedes Element von H von der Form ($\omega = \sigma^a \varrho\sigma \cdots \varrho\sigma\varrho^c =$)

$$\sigma^a (\varrho\sigma)^b \varrho^c$$

sein.

Da nun nach (3) beim Homomorphismus (9)

$$\varrho\sigma \rightarrow (0, 0, 1)(1, 0, 0) = (0, 1, 0)$$

gilt, so gilt dabei (nach (3))

$$\sigma^a (\varrho\sigma)^b \varrho^c \rightarrow (1, 0, 0)^a (0, 1, 0)^b (0, 0, 1)^c = (a, 0, 0)(0, b, 0)(0, 0, c) =$$
$$= (a, b, 0)(0, 0, c) = (a, b, c).$$

Wir haben gewonnen, daß durch (9) verschiedene Elemente von H auf solche von H' abgebildet werden. Somit ist H isomorph zu H'.

Nebenbei entstand auch der Beweis, daß (2) die eindeutige Darstellung aller Elemente von H ist.

Aus (1) folgt, daß ϱ eine Linkseinheit von H und auch von $R(H)$ ist. Da für das beliebige Element (2) von H

$$\sigma^a (\varrho\sigma)^b \varrho^c \varrho = \sigma^a (\varrho\sigma)^b \varrho^{c+1}$$

ist, so hat H und offenbar auch $R(H)$ kein Linkseinselement. Somit is der Satz bewiesen.

BEMERKUNG. Ein Element ε einer (multiplikativen) Struktur S mit der Eigenschaft $\varepsilon S = S\varepsilon = S$ werde eine Einheit von S genannt. Hat S eine

Einheit ε, so hat S auch sein Einselement. Wegen $\varepsilon S = S$ gibt es nämlich ein $\varepsilon'(\in S)$ mit $\varepsilon\varepsilon' = \varepsilon$, ferner läßt sich wegen $S\varepsilon = S$ jedes Element von S in der Form $\alpha\varepsilon$ $(\alpha \in S)$ annehmen. Da $(\alpha\varepsilon)\varepsilon' = \alpha(\varepsilon\varepsilon') = \alpha\varepsilon$ ist, so ist ε' ein Rechtseinselement von S. Da in S genau so die Existenz eines Linkseinselementes ε'' folgt, so ist $\varepsilon'(= \varepsilon''\varepsilon' = \varepsilon'')$ das Einselement von S.

(Eingegangen am 14. November 1959.)

KLASSIFIKATION DER DARSTELLUNGEN EINER GRUPPE

Von

F. LOONSTRA ('s Gravenhage, Holland)[1]

(*Vorgelegt von* L. RÉDEI)

Eine Gruppe B kann in mannigfacher Weise als homomorphes Bild dargestellt werden. Wenn G eine Gruppe, α eine homomorphe Abbildung von G auf B ist, so sprechen wir von einer *Darstellung* $G(\alpha)$ von B. Wir fragen, ob es möglich ist, eine Übersicht über alle möglichen Darstellungen $G(\alpha)$ von B zu bekommen. Die Untersuchung wird gestützt

1. auf eine Klassifikation der Darstellungen mit Hilfe des Begriffes der Verwandtschaft;

2. auf die Möglichkeit der Bildung eines Produktes von zwei Darstellungen;

3. auf die Möglichkeit einer Ordnungsrelation.

Wir bezeichnen den Kern der homomorphen Abbildung α von G auf B mit K_α.

1. Die Klassifikation der Darstellungen. Wir benutzten die folgende Eigenschaft von homomorphen Abbildungen: Induziert eine homomorphe Abbildung η der Gruppe G in die Gruppe G' einen Homomorphismus des Normalteilers $N \subseteqq G$ in den Normalteiler $N' \subseteqq G'$,

$$N\eta \subseteqq N', \quad G\eta \subseteqq G',$$

so induziert η auch einen Homomorphismus $\bar{\eta}$ der Faktorgruppe G/N in die Faktorgruppe G'/N' gemäß der Zuordnung

$$\bar{\eta}: \quad Ng \rightarrow N'g\eta \quad \text{für jedes } g \in G.$$

DEFINITION. Zwei Darstellungen $G(\alpha)$ und $G'(\alpha')$ sind *verwandt*, wenn verbindende Homomorphismen η und η' zwischen $G(\alpha)$ und $G'(\alpha')$ existieren:

1. $G\eta \subseteqq G', \quad G'\eta' \subseteqq G;$
2. $\alpha = \eta\alpha', \quad \alpha' = \eta'\alpha.$

[1] Nach meinem Vortrag „Über subdirekte Produkte" machte Herr WIELANDT eine Bemerkung über die Möglichkeit, das subdirekte Produkt beim Studium der Erweiterungen mit fester Faktorgruppe anzuwenden. Wiewohl Herr WIELANDT sich die Klassifikation dieser Darstellungen in anderer Richtung denkt, habe ich Herrn WIELANDT herzlichen Dank zu sagen.

Diese letzte Bedingung besagt, daß $K_\alpha \eta \subset K_{\alpha'}$, $K_{\alpha'} \eta' \subset K_\alpha$, und daß die Homomorphismen η und η' in B die identische Abbildung induzieren.

Beispiel. Ist $B = C_5$ die additive Gruppe der Restklassen mod 5, C die Gruppe der ganzen Zahlen, $C[x]$ die Gruppe der ganzzahligen Polynome in x, α die Abbildung von $n \in C$ auf den Rest von n nach Division durch 5 und β die Abbildung von $f(x) = a_n x^n + \cdots + a_1 x + a_0 \in C[x]$ auf den Rest von a_0 nach Division durch 5, so sieht man leicht, daß $C(\alpha)$ und $C[x](\beta)$ zwei verwandte Darstellungen von C_5 sind. Die Gruppe C besitzt überdies die Eigenschaft, daß — für jede homomorphe Abbildung α von C auf C_5 — die sämtlichen Darstellungen $C(\alpha)$ von C_5 untereinander verwandt sind. Die gegebene Definition der Verwandtschaft von Darstellungen ist eine Äquivalenzrelation, wie man leicht bestätigt. Bezeichnet man die identische Abbildung von B auf sich mit ε und ist α eine isomorphe Abbildung von B auf sich, so sind $B(\varepsilon)$ und $B(\alpha)$ zwei verwandte Darstellungen.

Satz I. *Die Darstellung $G(\alpha)$ ist genau dann zur Darstellung $B(\varepsilon)$ verwandt, wenn G nach dem Kern K_α zerfällt, d. h. wenn G dem Produkt $K_\alpha \cdot B'$ des Normalteilers K_α und einer Untergruppe $B' \simeq B$ mit $K_\alpha \cap B' = e$ gleich ist* (ε ist die identische Abbildung von B auf sich).

Beweis. Ist die Darstellung $G(\alpha)$ mit $B(\varepsilon)$ verwandt, so gibt es eine homomorphe Abbildung η' von B in G, die die Elemente von B invariant läßt, d. h. η' ist eine isomorphe Abbildung von B in G; G enthält also eine Untergruppe B' mit $G/K_\alpha \simeq B' \simeq B$, $B' \cap K_\alpha = e$, $G = K_\alpha \cdot B'$. Umgekehrt: es sei $G(\alpha)$ eine Darstelluug derart, daß G eine Untergruppe $B' \sim B$ enthält mit $G = K_\alpha \cdot B'$, $K_\alpha \cap B' = e$. Jedes Element $g \in G$ hat also die eindeutige Form $g = ab'$, $a \in K_\alpha$, $b' \in B'$. Die Abbildung $g = ab' \rightarrow (ab')\alpha = b$ ist eine homomorphe Abbildung von G auf B, die eine isomorphe Abbildung φ von B' auf B induziert und B invariant läßt. Ferner ist φ^{-1} eine isomorphe (also homomorphe) Abbildung von B in G, so daß B invariant ist. Es sind also $G(\alpha)$ und $B(\varepsilon)$ verwandte Darstellungen.

Satz II. *Ist B eine freie Gruppe, so sind alle Darstellungen $G(\alpha)$ von B verwandt und umgekehrt.*

Beweis. Ist $K_\alpha g_i = b_i$ ein freies erzeugendes System von G/K_α, so erzeugen die Elemente g_i eine freie Untergruppe $F \subset G$ mit $F \cap K_\alpha = e$, also $G = K_\alpha \cdot F$, $K_\alpha \cap F = e$, $F \simeq B$, also die Darstellungen $G(\alpha)$ und B sind verwandt. Sind umgekehrt alle Darstellungen $G(\alpha)$ mit $B(\varepsilon)$ verwandt, so zerfällt auch jede Darstellung $G(\alpha)$, in der G frei ist. Es ist $G = K_\alpha \cdot B'$, $K_\alpha \cap B' = e$, $B' \simeq B$, also B' und B sind frei.

Satz III. *Zwei Darstellungen $G(\alpha)$ und $G'(\alpha')$ sind verwandt, wenn G und G' frei sind.*

Beweis. Zum Beweise benutzen wir folgendes: ist $\bar{\gamma}$ eine homomorphe Abbildung der freien Gruppe F in die Faktorgruppe G/A, so wird $\bar{\gamma}$ durch eine homomorphe Abbildung γ von F in G induziert. In der Tat, die Abbildung

$$\bar{\gamma}:\quad f \to Ag(f) \in G/A \quad (\text{für jedes } f \in F)$$

bestimmt mit einem freien Erzeugendensystem $S=(s)$ von F durch

$$\gamma:\quad s \to g(s) \quad \text{für jedes} \quad s \in S$$

einen Homomorphismus γ von F in G, der den Homomorphismus

$$\bar{\gamma}:\quad s \to g(s) \to Ag(s) \quad \text{für jedes} \quad s \in S$$

der Gruppe F in G/A induziert.

Sind nun $G(\alpha)$ und $G'(\alpha')$ Darstellungen, so gibt es eine isomorphe Abbildung

$$\varphi:\quad G/K_\alpha \leftrightarrow G'/K_{\alpha'}$$

derart, daß entsprechende Elemente $\bar{g} \leftrightarrow \bar{g}'$ mit demselben Element $b \in B$ übereinstimmen. Wir schreiben $(G/K_\alpha)\varphi = G'/K_{\alpha'}$ und beachten, daß φ die identische Abbildung in B induziert. Setzen wir nun voraus, daß $G(\alpha)$ und $G'(\alpha')$ Darstellungen sind mit freien G, G', so definieren wir

$$\bar{\gamma}:\quad g \to (gK_\alpha)\varphi \in G'/K_{\alpha'},$$
$$\bar{\gamma}':\quad g' \to (g'K_{\alpha'})\varphi^{-1} \in G/K_\alpha;$$

das sind homomorphe Abbildungen von G auf $G'/K_{\alpha'}$ (bzw. von G' auf G/K_α). Dann gibt es also auch eine homomorphe Abbildung γ (bzw. γ') von G in G', welche K_α in $K_{\alpha'}$ abbildet (bzw. von G' in G usw.), so daß in B die identische Abbildung induziert wird. Also sind die Darstellungen $G(\alpha)$ und $G'(\alpha')$ verwandt.

2. Das Produkt von Darstellungen.

Definition. Wenn $G(\alpha)$ und $G'(\alpha')$ zwei Darstellungen sind, so ist das Produkt dieser Darstellungen das sogenannte meromorphe subdirekte Produkt $P(\beta) = G(\alpha) \underset{\sim}{\times} G'(\alpha')$; d. h. $P(\beta)$ ist die Gruppe der Paare (g, g'), $g \in G$, $g' \in G'$ mit der Eigenschaft

$$(g, g')\beta = g\alpha = g'\alpha'.$$

Dieses Produkt ist eine Darstellung von B, denn $G(\alpha) \underset{\sim}{\times} G'(\alpha')$ enthält den Normalteiler $K_\alpha \times K_{\alpha'}$ und es ist

$$[G(K_\alpha) \underset{\sim}{\times} G'(K'_{\alpha'})]/[K_\alpha \times K_{\alpha'}] \cong G/K_\alpha \cong G'/K_{\alpha'} \cong B.$$

Satz IV. *Sind die Darstellungen $G(\alpha)$ und $G'(\alpha')$ verwandt, so sind auch $G(\alpha) \underset{\sim}{\times} \bar{G}(\bar{\alpha})$ und $G'(\alpha') \underset{\sim}{\times} \bar{G}(\bar{\alpha})$ für jede Darstellung $\bar{G}(\bar{\alpha})$ verwandt, d. h. die Klasse des Produktes wird durch die Klassen der Faktoren bestimmt.*

Beweis. Es seien η, η' die verbindenden Homomorphismen zwischen G und G'; $G(\alpha) \underset{\sim}{\times} \overline{G}(\overline{\alpha})$ enthält genau die Elemente $(g', \overline{g})$ mit $g\alpha = \overline{g}\,\overline{\alpha}$ und $G'(\alpha') \underset{\sim}{\times} \overline{G}(\overline{\alpha})$ enthält genau die Elemente $(g', \overline{g})$ mit $g'\alpha' = \overline{g}\,\overline{\alpha}$. Wir definieren nun die verbindenden Homomorphismen ζ, ζ' folgenderweise:

$$(g, \overline{g})\zeta = (g\eta, \overline{g}), \quad (g', \overline{g})\zeta' = (g'\eta', \overline{g})$$

für jedes Paar $(g, \overline{g}) \in G \underset{\sim}{\times} \overline{G}$, $(g', \overline{g}) \in G' \underset{\sim}{\times} \overline{G}$. Die Abbildung ζ bildet in der Tat $(g, \overline{g})$ auf ein Element von $G' \underset{\sim}{\times} \overline{G}$ ab, denn es ist $g\alpha = (g\eta)\alpha'$, $(a, \overline{a})\zeta = (a\alpha, \overline{a})$, daher $(K_\alpha \times K_{\overline{\alpha}})\zeta \subseteq K_{\alpha'} \times K_{\overline{\alpha}}$ und ζ induziert in B die identische Abbildung, und analoges gilt für die Abbildung ζ'. Sind also $G(\alpha)$ und $G'(\alpha')$ verwandte Darstellungen, dann sind es auch $G(\alpha) \underset{\sim}{\times} \overline{G}(\overline{\alpha})$ und $G'(\alpha') \underset{\sim}{\times} \overline{G}(\overline{\alpha})$.

3. Ordnung der Klassen.

Definition. *Wenn $G(\alpha)$ und $G'(\alpha')$ zwei Darstellungen sind, so sei*

$$G(\alpha) \leqq G'(\alpha'),$$

wenn es einen Homomorphismus η von G in G' gibt, so daß $G\eta \subseteq G'$ und $\alpha = \eta\alpha'$, d. h. η induziert die identische Abbildung in B. Diese Ordnung der Darstellungen ist eine Klassenordnung:

Satz V. *Wenn $G(\alpha) \leqq G'(\alpha')$ und $G(\alpha)$ mit $\overline{G}(\overline{\alpha})$ verwandt ist, so gilt auch $\overline{G}(\overline{\alpha}) \leqq G'(\alpha')$ (und analog: wenn $G'(\alpha')$ mit $\overline{G}'(\overline{\alpha}')$ verwandt ist usw.).*

Beweis. Es sei ζ eine verbindende homomorphe Abbildung von $\overline{G}$ in G mit $\overline{G}\zeta \subseteq G$, $\overline{\alpha} = \zeta\alpha$; aus $G(\alpha) \leqq G'(\alpha')$ folgt die Existenz einer Abbildung η mit $G\eta \subseteq G'$, $\alpha = \eta\alpha'$. Die homomorphe Abbildung $\xi = \zeta\eta$ hat die Eigenschaften

$$(\overline{G})\xi = (\overline{G})\zeta\eta \subseteq G\eta \subseteq G', \quad \zeta\eta\alpha' = \overline{\alpha},$$

d. h. $\xi = \zeta\eta$ induziert in B die identische Abbildung. Daraus folgt $\overline{G}(\overline{\alpha}) \leqq$ $\leqq G'(\alpha')$. Die Ordnungsrelation ist also eine Klasseneigenschaft. Es gilt darum auch $G(\alpha) \leqq G(\alpha)$.

Satz VI. *Die Ordnungsrelation ist transitiv: aus $G(\alpha) \leqq G'(\alpha')$ und $G'(\alpha') \leqq G''(\alpha'')$ folgt $G(\alpha) \leqq G''(\alpha'')$.*

Beweis. Klar!

Satz VII. *Aus $G(\alpha) \leqq G'(\alpha')$, $G'(\alpha') \leqq G(\alpha)$ folgt, daß $G(\alpha)$ und $G'(\alpha')$ verwandt sind.*

Beispiel. Ist wieder $B = C_5$, so ist $C(\alpha)$ eine Darstellung (siehe S. 224) und es gilt $C(\alpha) \leqq C_5$, aber nicht umgekehrt; man kann n auf die Restklasse n mod 5 abbilden, und diese Abbildung genügt allen Forderungen. Es gibt aber keine derartige Abbildung von C_5 in $C(\alpha)$.

SATZ VIII. *Immer gilt $G(\alpha) \leqq B(\varepsilon)$, d. h. die Klasse von B ist die größte gemäß der Ordnungsrelation.*

BEWEIS. Die natürliche homomorphe Abbildung α von G auf B genügt allen Eigenschaften.

SATZ IX. *Es sei M ein System von Darstellungen $G_i(\alpha_i)$. Es folgt die Existenz einer Darstellung $P(\alpha)$ mit den Eigenschaften:*

1. *$P(\alpha) \leqq G_i(\alpha_i)$ für jede $G_i(\alpha_i) \in M$;*
2. *aus $H(\beta) \leqq G_i(\alpha_i)$ für jede $G_i(\alpha_i) \in M$ folgt $H(\beta) \leqq P(\alpha)$. Das bedeutet: für jedes bestimmte System M von Darstellungen existiert die größte untere Grenze (Produkt $= P$).*

BEWEIS. Man bildet das meromorphe subdirekte Produkt $P = \underset{\widetilde{i}}{\times}\, G_i(\alpha_i)$; P ist die Gruppe aller Systeme (g_i) mit der Eigenschaft $g_i\alpha_i = g_j\alpha_j$ $(i \neq j)$. $P(\alpha)$ ist eine Darstellung, denn $P(\alpha)$ enthält das direkte Produkt $K_\alpha = \underset{i}{\times}\, K_{\alpha_i}$ und es gilt

$$P/K_\alpha \simeq G_i/K_{\alpha_i} \simeq B, \quad \text{also} \quad p\alpha = (g_i)\alpha = g_j\alpha_j.$$

Wir bilden nun ab

$$p = c(g_i) \to (g_i)\eta_m = g_m;$$

das ist eine homomorphe Abbildung von P auf G_k, so daß K_α auf K_{α_m} abgebildet wird und B invariant bleibt, also $P(\alpha) \leqq G_i(\alpha_i)$ für jedes i. Ist umgekehrt

$$H(\beta) \leqq G_i(\alpha_i) \quad \text{für jedes} \quad i,$$

so gibt es homomorphe Abbildungen η_i von H in G_i, also

$$H\eta_i \subseteqq G_i, \quad \beta = \eta_i\alpha_i.$$

Die Abbildung

$$\eta: \quad h \to h\eta = (h\eta_i) \quad \text{für jedes} \quad h \in H$$

ist eine Abbildung von H in $P(\alpha) = \underset{\widetilde{i}}{\times}\, G_i/K_{\alpha_i}$ und genügt den drei Eigenschaften, also $H(\beta) \leqq P(\alpha)$.

FOLGERUNG. Das Produkt $G(\alpha) \underset{\sim}{\times} G'(\alpha')$ zweier Darstellungen ist die größte untere Grenze der beiden Darstellungen $G(\alpha)$ und $G'(\alpha')$. Ist $G(\alpha) \leqq$ $\leqq G'(\alpha')$, so sind die Darstellungen $G(\alpha) \underset{\sim}{\times} G'(\alpha')$ und $G(\alpha)$ verwandt! Wenn man nun die Gesamtheit der Klassen von Darstellungen betrachtet, so kann man folgendes sagen: Diese ist teilweise geordnet; jedes System M von Klassen hat eine größte untere Grenze und es gibt außerdem eine größte Klasse (die Klasse von $B(\varepsilon)$). Dann hat auch jedes bestimmte System M von Klassen von Darstellungen eine kleinste obere Grenze. Wir beweisen nun noch 3 von:

Satz X. *Ist M ein bestimmtes System von Darstellungen $G_i(\alpha_i)$ von B, so gibt es eine Darstellung $S(\gamma)$ (= Summe) mit den Eigenschaften*

1. $G_i(\alpha_i) \leqq S(\gamma)$ *für jede* $G_i(\alpha_i) \in M$;
2. *wenn* $G_i(\alpha_i) \leqq H(\delta)$ *für jede* $G_i(\alpha_i) \in M$, *so gilt* $S(\gamma) \leqq H(\delta)$;
3. *das freie Produkt* $S = \underset{i}{*} G_i(\alpha_i)$ *der Darstellungen* $G_i(\alpha_i)$ *genügt allen Bedingungen der Summe.*

Beweis. Jedes Element $(\neq e)$ von S ist ein Wort

$$s = g_{i_1} g_{i_2} \dots g_{i_n}, \quad n > 0, \quad g_{i_\nu} \in G_{i_\nu}, \quad i_{\nu-1} \neq i_\nu \quad \text{für} \quad 2 \leqq \nu \leqq n.$$

Bildet man ab.

$$s = g_{i_1} g_{i_2} \dots g_{i_n} \to s(\gamma) = (g_{i_1} \alpha_{i_1}) \dots (g_{i_n} \alpha_{i_n}) = b_{i_1} b_{i_2} \dots b_{i_n},$$

so hat man eine homomorphe Abbildung von S auf B. $S(\gamma)$ ist also eine Darstellung. K_γ enthält den von den K_{α_i} in S erzeugten Normalteiler A. Es ist $G_i(\alpha_i) \leqq S(\gamma)$, denn die identische Abbildung von G_i in S bildet K_{α_i} in K_γ ab und läßt B invariant.

Es sei umgekehrt $H(\delta)$ eine Darstellung von B mit der Eigenschaft, daß für jede Darstellung $G_i(\alpha_i)$ des Systems M $G_i(\alpha_i) \leqq H(\delta)$ gilt, d. h., es gibt für jedes i eine homomorphe Abbildung η_i von G_i in H, so daß $\alpha_i = \eta_i \delta$. Dann ist aber auch $S(\gamma) = \underset{i}{*} G_i(\alpha_i) \leqq H(\delta)$. Definiert man nämlich die homomorphe Abbildung η von $S(\gamma) = \underset{i}{*} G_i(\alpha_i)$ in $H(\delta)$ durch

$$s = g_{i_1} g_{i_2} \dots g_{i_n} \to (g_{i_1} \eta_{i_1}) \cdot (g_{i_2} \eta_{i_2}) \dots (g_{i_n} \eta_{i_n}),$$

so ist η eine homomorphe Abbildung von S in H. $s = g_{i_1} g_{i_2} \dots g_{i_n} \in K_\gamma$ bedeutet, daß

$$(g_{i_1} \alpha_{i_1}) \dots (g_{i_n} \alpha_{i_n}) = b_{i_1} \cdots b_{i_n} = e.$$

Aber die homomorphen Abbildungen η_{i_k} lassen B invariant; es ist daher auch das Produkt der natürlichen homomorphen Bilder von $g_{i_k} \eta_{i_k}$ $(k = 1, 2, \dots, n)$ wieder die Identität, also $K_\gamma \eta \subseteqq K_\delta$, und B bleibt invariant. Somit ist S die Summe der Darstellungen $G_i(\alpha_i)$ des Systems M.

Satz XI. *Folgende Aussagen sind gleichwertig:*

1. *Für die Darstellungen* $G(\alpha)$ *und* $\bar{G}(\bar{\alpha})$ *gilt* $G(\alpha) \leqq \bar{G}(\bar{\alpha})$;
2. $\bar{G}(\bar{\alpha})$ *ist homomorphes Bild einer Darstellung, die G enthält;*
3. $\bar{G}(\bar{\alpha})$ *ist homomorphes Bild einer Darstellung, die eine mit G isomorphe Retrakte enthält.*

($\bar{G}(\bar{\alpha})$ sei homomorphes Bild einer Darstellung $F(\beta)$, dies bedeutet hier: es gibt eine homomorphe Abbildung von $F(\beta)$ auf $\bar{G}(\bar{\alpha})$, die K_β in $K_{\bar{\alpha}}$ abbildet und in B die identische Abbildung induziert. Daß die Darstellung $F(\beta)$ die

Darstellung $G(\alpha)$ enthält, bedeutet, daß es eine isomorphe Abbildung von G in F gibt, die K_α in K_β abbildet und B invariant läßt.)

Beweis. Aus 2 folgt 1, denn $G(\alpha) \leqq F(\beta) \leqq \bar{G}(\bar{\alpha})$, also $G(\alpha) \leqq \bar{G}(\bar{\alpha})$.

Aus 1 folgt 3: es gilt $G(\alpha) \leqq \bar{G}(\alpha)$, also gibt es eine homomorphe Abbildung von G in $\bar{G}$ mit

$$G\eta \subseteqq \bar{G}, \quad \alpha = \eta\bar{\alpha}.$$

Wir betrachten nun die Gruppe $F = G(\alpha) \underset{\sim}{\times} \bar{G}(\bar{\alpha})$ der Paare $(g, \bar{g})$ mit $g\alpha = \bar{g}\bar{\alpha}$. Die Abbildung

$$g \to (g, g\eta)$$

ist eine Abbildung von G in das Produkt $G(\alpha) \underset{\sim}{\times} \bar{G}(\bar{\alpha})$, denn g und $g\eta$ entsprechen demselben Element $b \in B$. Nun ist die Gruppe $\bar{F}$ der Paare $(g, g\eta)$ eine Untergruppe von F und es ist $\bar{F} \sim G(\alpha)$. Es ist also $F(\beta)$ mit $\beta(g, g\eta) = g\alpha$ eine Darstellung mit dem Kern bestehend aus den Elementen $(a, a\eta), a \in K_\alpha$. $G(\alpha)$ ist mit $\bar{F}(\beta)$ verwandt; die Abbildung η von $g \in G(\alpha)$ auf $(g, g\eta)$ bildet $a \in K_\alpha$ auf $(a, a\eta)$ ab und läßt B invariant. Die Abbildung $(g, g\eta)\bar{\eta} = g$ ist eine homomorphe Abbildung von $\bar{F}(\beta)$ in $G(\alpha)$ und bildet $(a, a\eta)$ auf a ab und induziert in B die identische Abbildung.

Wir betrachten nun die Gruppe $F = F(\alpha) \underset{\sim}{\times} \bar{G}(\bar{\alpha})$ mit den Elementen $(g, \bar{g})$ und außerdem die Abbildung

$$(g, \bar{g}) \to (g, g\eta).$$

Diese Abbildung ist eine idempotente endomorphe Abbildung und sie läßt $\bar{F}$ elementsweise invariant. Es ist also $\bar{F}$ eine Rektrakte von $F = G(\alpha) \underset{\sim}{\times} \bar{G}(\bar{\alpha})$. Die Abbildung $(g, \bar{g}) \to \bar{g}$ für jedes Paar von Elementen von F ist eine homomorphe Abbildung von $G(\alpha) \underset{\sim}{\times} \bar{G}(\bar{\alpha})$ auf $\bar{G}$, die $K_\alpha \times K_{\bar{\alpha}}$ auf $K_{\bar{\alpha}}$ abbildet und B invariant läßt. Deshalb ist $\bar{G}(\bar{\alpha})$ ein homomorphes Bild der Darstellung F und F enthält eine mit G isomorphe Retrakte $\bar{F}$.

Da 2 aus 3 folgt, haben wir die Gleichwertigkeit der drei Aussagen bewiesen.

(Eingegangen am 14. November 1959.)

SOME INTERSECTION PROPERTIES OF RANDOM WALK PATHS*

By

P. ERDŐS (Budapest), corresponding member of the Academy,
and S. J. TAYLOR (Birmingham)

1. Introduction. We are concerned with properties of the infinite symmetric random walk over the lattice of points with integer coordinates in Euclidean space of d dimensions. (For a precise definition see [4].) We consider two problems.

PROBLEM A. *Suppose $\Pi^{(d)}(a, b)$ denotes the set of points $S_d(n)$ $(a \leqq n \leqq b)$ of the random walk in d-space; suppose $f(n)$ is an increasing function of n. What are the conditions on the rate of increase of $f(n)$ which are necessary and sufficient to ensure that the sets*

$$\Pi^{(d)}(0, n), \quad \Pi^{(d)}(n+f(n), \infty)$$

have points in common for infinitely many values of n with probability 1?

We complete the solution of this problem in Section 3. Clearly, there is no problem for $d=1$ or 2. The solution takes a different form in the cases $d=3$, $d=4$, and $d \geqq 5$. For example, if $d=4$, an interesting consequence of the result is that, with probability 1, there are infinitely many n for which $\Pi^4(0, n)$ and $\Pi^4(2n, \infty)$ have a point in common. This in turn implies that any two independent random walks in 4-space have infinitely many points in common. This at first surprised us, because two independent Brownian motion paths in 4-space have no points in common with probability 1 (this follows from the result of [2]). The explanation is as follows: with probability 1, two independent Brownian paths in 4-space approach arbitrarily close to each other for arbitrarily large values of t; thus they have infinitely many near misses, but fail to intersect because the fine structure of the paths is not sufficiently dense (in fact, the paths have zero 2-dimensional measure, see [5]). It can be shown similarly that, with probability 1, 3 independent random walks in 3-space have infinitely many points in common, while 3 Brownian motion paths have no common points (this last follows from the result of [3]).

* This paper and the paper [4] were written while P. ERDŐS was visiting the University of Birmingham.

PROBLEM B. *A point $S_d(n)$ of a random walk path is called 'good' if there are no points common to $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+1, \infty)$. For $d=1$ or 2 there are no good points with probability 1. For $d \geqq 3$, the result of Pólya implies that there must be some good points: how many are there?*

We prove, in Section 4, that there are absolute constants τ_d $(d \geqq 5)$ such that, with probability 1, a random walk in d-space has good points at a subsequence of density τ_d. For $d=3$ or 4 the subsequence of good points has density zero with probability 1. We obtain in these cases asymptotic bounds for the number of good points.

We start, in Section 2, by obtaining some preliminary results, and collecting results which are already known but are needed in the sequel.

2. Preliminary results. If E is a condition on the random walk path, we write $\mathbf{P}(E)$ for the probability that the condition is satisfied. If $E_1, E_2, \ldots, E_k, \ldots$ is a sequence of conditions, we write

$$\mathbf{P}\{E_k \ i.o.\}$$

for the probability that the path satisfies infinitely many of the conditions E_k.

$\mathcal{E}\{Q\}$, $\sigma^2\{Q\}$ denote the mean and variance of a random variable Q.

$[x]$ denotes the largest integer not greater than the real number x.

ε will always denote a positive number.

$c_1, c_2, \ldots, c_{58}$ will denote suitable finite positive real constants.

If X is a vector in d-space, $|X|$ denotes the distance from X to the origin.

For paths in d-space, $\gamma_d(n)$ denotes the probability that in the first $n-1$ steps, the path does not return to the origin. It is proved in [1] that, for $d \geqq 3$, there are positive constants γ_d such that

$$\gamma_d < \gamma_d(n) < \gamma_d + O(n^{1-d/2}). \tag{2.1}$$

$S_d(n)$ denotes the position at the n^{th} step of a random walk in d-space. If L is any lattice point in d-space,

$$u_d(L, n) = \mathbf{P}\{S_d(n) = L\}.$$

Clearly, all points can be reached either only in an even number of steps or only in an odd number of steps, We need the following easy estimates for $u_d(L, n)$. Suppose L is a point which can be reached in an even number of steps, and $|L| = \varrho$. Then

(i) if $\varrho = 0$, we have

$$u_d(0, 2m) = 2\left(\frac{d}{4m\pi}\right)^{d/2} + O\left(\frac{1}{m^{1+d/2}}\right); \tag{2.2}$$

(ii) if $m > \varrho^2$,

$$u_d(L, 2m) = 2\left(\frac{d}{4m\pi}\right)^{d/2}\left[1 + O\left(\frac{\varrho^2}{m}\right)\right]; \tag{2.3}$$

(iii) if $m > \frac{1}{20}\varrho^2$, there are constants c_1, c_2 with

$$\frac{c_1}{m^{d/2}} > u_d(L, 2m) > \frac{c_2}{m^{d/2}}; \tag{2.4}$$

(iv) if $m < \varrho^2, d \geqq 2$,

$$u_d(L, 2m) \leqq 2\left(\frac{d}{4m\pi}\right)^{d/2} e^{-\varrho^2/2m}\left(1 + O\left(\frac{1}{m}\right)\right). \tag{2.5}$$

The same asymptotic formulae are valid for $u_d(L, 2m+1)$ in the case where L can be reached in an odd number of steps.

We now need estimates for entering a point L at least once.

LEMMA 1. *Suppose L is a lattice point in d-space ($d \geqq 3$), with $|L| = \varrho > 0$, and $v_d(L)$ is the probability that an infinite random walk starting from O will enter L at least once. Then there are finite positive constants f_d, g_d such that*

$$\frac{f_d}{\varrho^{d-2}} < v_d(L) < \frac{g_d}{\varrho^{d-2}} \qquad (d = 3, 4, \ldots).$$

PROOF. We may clearly assume that $\varrho > 100$. Considering the last time of passage through L, we have

$$v_d(L) \geqq \gamma_d \sum_{n=1}^{\infty} u_d(L, n),$$

since $\gamma_d u_d(L, n)$ is the probability that the path is at L at the n^{th} step and does not return again to L and these events are mutually exclusive. Using (2. 4) we have

$$v_d(L) \geqq \gamma_d \sum_{m > \frac{1}{20}\varrho^2}^{\infty} \frac{c_2}{m^{d/2}} \geqq \frac{f_d}{\varrho^{d-2}} \quad \text{for a suitable } f_d > 0.$$

In the other direction, it is clear that

$$v_d(L) \leqq \sum_{n=1}^{\infty} u_d(L, n) = \sum_{n \leqq \frac{1}{20}\varrho^2} u_d(L, n) + \sum_{n > \frac{1}{20}\varrho^2} u_d(L, n) \leqq \frac{g_d}{\varrho^{d-2}}$$

for a suitable $g_d > 0$,

using (2. 5) and (2. 4).

This completes the proof of the lemma.

A modified form of the same proof suffices to prove

LEMMA 2. *Suppose L is a lattice point in d-space ($d \geqq 3$) with $|L| = \varrho > 0$ and $w_d(L, n)$ is the probability that a random walk of n steps starting from O will enter L at least once. Then there are finite constants h_d such that if $n > \frac{1}{5}\varrho^2$, then*

$$w_d(L, n) > \frac{h_d}{\varrho^{d-2}} \qquad (d = 3, 4, \ldots).$$

For $d \geqq 3$, we know that random walks wander off to infinity. We need estimates for the probability that they are not too far from O at some time in a given range of t.

LEMMA 3. *For every integer $N \geqq 0$ and real number $r > 0$, put*

$$Q_d(r, N) = \mathbf{P}\{|S_d(n)| \leqq r \text{ for some } n \geqq N\},$$

then we have for $d \geqq 3$, $N > r^2$

$$\frac{1}{2} e_d \left(\frac{r}{\sqrt{N}}\right)^{d-2} \leqq Q_d(r, N) \leqq e_d \left(\frac{r}{\sqrt{N}}\right)^{d-2}$$

for suitable $e_d > 0$.

This gives the probability of being within a distance r of the origin at some time after the N^{th} step. The corresponding result for Brownian motion is proved in [1]; the random walk result follows because of the relationship between random walk and Brownian motion.

A proof is also given in [1] for

LEMMA 4. *For every integer $N \geqq 0$ and real number $r > 0$, put*

$$P_d(r, N) = \mathbf{P}\{S_d(n) \leqq r \text{ for some } N \leqq n \leqq 4N\};$$

then for $d \geqq 3$, $P_d(r, N) > \frac{1}{10} e_d \left(\frac{r}{\sqrt{N}}\right)^{d-2}$.

For a random walk in d-space, we put $\varrho_d(n) = |S_d(n)|$. In [4] we briefly studied the average behaviour of $\varrho_d(n)$ and in doing so proved

LEMMA 5. *If $\varrho_4(n)$ is the distance from the origin at the n^{th} step of a random walk in 4 dimensions, then there exists a constant $c_8 > 0$ such that*

$$\mathcal{E}\left(\frac{1}{1 + \{\varrho_4(n)\}^2}\right) = \frac{c_8}{n}(1 + o(1)).$$

Similarly one can prove

LEMMA 6. *If $\varrho_3(n)$ is the distance from the origin at the n^{th} step of a random walk in 3-space, then there exist constants c_3', c_3'' such that*

$$\frac{c_3'}{\sqrt{n}} < \mathcal{E}\left\{\frac{1}{1+\varrho_3(n)}\right\} < \frac{c_3''}{\sqrt{n}}.$$

3. Solution of Problem 1. In order to estimate the probability of an intersection between $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+f(n), \infty)$ we need first of all to estimate the probability of at least one intersection between two independent random walks starting at different points. It is critical to suppose that if the starting points are separated by a distance ϱ, then one of the paths takes approximately ϱ^2 steps, while the other either takes ϱ^2 steps or is infinite.

LEMMA 7. *Suppose $\Pi_1^{(d)}(0, n)$, $\Pi_2^{(d)}(0, n)$ are independent random walk paths of n steps in d-space ($d \geqq 3$), the first starting at the origin and the second at P where $|P| = \varrho$ and $\frac{1}{2}\varrho^2 < n < 2\varrho^2$; then there are constants $c_4, c_5, c_6^{(d)}$ such that*

(i) $\mathbf{P}\{\Pi_1^{(3)}(0, n)$ *intersects* $\Pi_2^{(3)}(0, n)\} > c_4$,

(ii) $\mathbf{P}\{\Pi_1^{(4)}(0, n)$ *intersects* $\Pi_2^{(4)}(0, n)\} > \dfrac{c_5}{\log n}$,

(iii) $\mathbf{P}\{\Pi_1^{(d)}(0, n)$ *intersects* $\Pi_2^{(d)}(0, n)\} > \dfrac{c_6^{(d)}}{n^{(d-4)/2}}$ *for* $d = 5, 6, \ldots$.

PROOF OF (i). In Euclidean space of 3 dimensions consider a cube b with centre at the origin and side $\frac{1}{2}\varrho$. Let η be a large positive number and

$$t_3 = [\eta n^{1/6}]. \tag{3.1}$$

Let $\mathfrak{L}$ be the subset of the lattice points with integer coordinates obtained by taking points all of whose coordinates are multiples of t_3 which lie in b but are not within $\frac{1}{4}\varrho$ of the origin. The number $\nu(\mathfrak{L})$ of points of $\mathfrak{L}$ clearly satisfies

$$\nu(\mathfrak{L}) \geqq c_7 \frac{n}{\eta^3} \tag{3.2}$$

and the distance of any point L in $\mathfrak{L}$ from each of O, P lies between $\frac{1}{4}\varrho$ and 2ϱ.

Let $p_1(L)$ denote the probability that $\Pi_1^{(3)}(0, n)$ passes through L, and $p_2(L)$ the probability that $\Pi_2^{(3)}(0, n)$ passes through L. Since the paths are independent, the probability that L lies on both paths is $p_1(L)p_2(L)$. Using Lemmas 1 and 2 we have immediately that

$$\frac{4g_3}{\varrho} \geqq \begin{Bmatrix} p_1(L) \\ p_2(L) \end{Bmatrix} \geqq \frac{h_3}{2\varrho} \tag{3.3}$$

so that

$$p_1(L)p_2(L) \geqq \frac{c_8}{n}. \tag{3.4}$$

If interference could be neglected, (3. 2) and (3. 4) would be sufficient to obtain the desired result. We now show that provided η is chosen large enough, the interference is small. Let $p_1(L, M)$, $p_2(L, M)$, respectively, be the probabilities that $\Pi_1^{(3)}(0, n)$, $\Pi_2^{(3)}(0, n)$ pass through both L and M. Clearly, it is sufficient to show that

$$\sum_{L \in \mathfrak{L}} p_1(L)p_2(L) - \sum_{L, M \in \mathfrak{L}}\sum p_1(L, M)p_2(L, M) > c_4. \tag{3.5}$$

In the notation of Lemma 1, $v_3(L)$ is the probability that an infinite path from the origin will pass through L. It is clear that

$$p_1(L, M) \leqq v_3(L-M)(p_1(L)+p_1(M)),$$
$$p_2(L, M) \leqq v_3(L-M)(p_2(L)+p_2(M)).$$

By (3. 3),

$$\sum p_1(L, M)p_2(L, M) \leqq \left(\frac{8g_3}{\varrho}\right)^2 \sum [v_3(L-M)]^2 < \frac{c_9}{n} \sum [v_3(L-M)]^2. \tag{3.6}$$

For fixed $L \in \mathfrak{L}$, let $\nu_r(L)$ be the number of points of $\mathfrak{L}$ whose distance from L lies between 2^r and 2^{r+1} $(r = 1, 2, \ldots)$. It is clear that

$$\nu_r(L) \leqq c_{10} \frac{2^{3r}}{\eta^3 n^{1/2}}. \tag{3.7}$$

For a point M whose distance from L is at least 2^r we know that

$$v_3(L-M) \leqq \frac{g_d}{2^r}.$$

Hence

$$\sum_{M \in \mathfrak{L}} v_3(L-M)^2 \leqq \sum_{1 \leqq 2^r \leqq \varrho} \nu_r(L) \frac{g_d^2}{2^{2r}} < \frac{c_{11}}{\eta^3},$$

using (3. 7). By (3. 6), it follows that, for each L in $\mathfrak{L}$,

$$\sum_{M \in \mathfrak{L}} p_1(L, M)p_2(L, M) < \frac{c_9}{n} \frac{c_{11}}{\eta^3} < \frac{1}{2} p_1(L)p_2(L),$$

by (3.3), provided η is chosen large enough. This together with (3.2) and (3.4) is clearly sufficient to establish (3.5).

PROOF OF (ii). A very similar argument will work using for $\mathfrak{L}$ only points whose co-ordinates are all divisible by

$$t_4 = [\eta (\log n)^{1/4}]$$

for large enough η.

PROOF OF (iii). Again the same argument works using this time points whose co-ordinates are all divisible by t_d where t_d is a sufficiently large fixed integer.

The simplest way of obtaining an upper bound corresponding to Lemma 7 (ii) seems to be a calculation of the expectation of the number of points common to the two paths by two different methods. We first obtain a lower bound for this expectation in the case where both paths start at the origin.

LEMMA 8. *Suppose $\Pi_1^{(d)}(0, n)$, $\Pi_2^{(d)}(0, n)$ are independent random walks of n steps in d-space ($d = 3$ or 4), both starting from the origin. Let $D^{(d)}(n)$ denote the number of points common to the two paths. Then there are constants c_{12}, c_{13} such that*

(i) $\mathcal{E}\{D^{(3)}(n)\} > c_{12} n^{1/2}$,

(ii) $\mathcal{E}\{D^{(4)}(n)\} > c_{13} \log n$.

PROOF OF (i). Consider the points P in 3-space with integer co-ordinates whose distance from O is less than $\sqrt{n}$. It is clear that if $v_3(P)$ is the probability that a path of n steps will enter P,

$$\mathcal{E}\{D^{(3)}(n)\} \geqq \sum_{1 \leqq |P| < n^{1/2}} [v_3(P)]^2 \geqq \sum_{1 \leqq 2^r < n^{1/2}} \; \sum_{2^{r-1} \leqq |P| < 2^r} [v_3(P)]^2.$$

If $2^r > |P|$, we have, by Lemma 2,

$$v_3(P) \geqq \frac{c_{14}}{|P|} > \frac{c_{14}}{2^r}.$$

As the number of points P with $2^{r-1} \leqq |P| < 2^r$ is at least $c_{15} 2^{3r}$, it follows immediately that

$$\mathcal{E}\{D^{(3)}(n)\} \geqq c_{12} n^{1/2}.$$

PROOF OF (ii). The same method works. As above,

$$\mathcal{E}\{D^{(4)}(n)\} \geqq \sum_{1 \leqq 2^r < n^{1/2}} \; \sum_{2^{r-1} \leqq |P| < 2^r} [V_4(P)]^2 \geqq c_{16} \sum_{r=0}^{[\frac{1}{2}\log_2 n]} 1 > c_{13} \log n.$$

REMARK 1. It is clear that there is a constant c_{17} such that for any n and $d \geqq 5$

$$1 \leqq \mathcal{E}\{D^{(d)}(n)\} < c_{17}.$$

REMARK 2. By carrying out the computations more carefully it is possible to prove that

$$\mathcal{E}\{D^{(3)}(n)\} = c_{18} n^{1/2}(1+o(1)),$$

$$\mathcal{E}\{D^{(4)}(n)\} = c_{19} \log n(1+o(1)),$$

$$\mathcal{E}\{D^{(d)}(n)\} = c_{20}^{(d)}(1+o(1)) \qquad (d \geqq 5),$$

for suitable constants $c_{18}, c_{19}, c_{20}^{(d)}$. We do not prove these as they are not required in the sequel.

Though we do not require it, we prove the following lemma to complete our results:

LEMMA 9. *Suppose $\Pi_1^{(d)}(0,\infty)$, $\Pi_2^{(d)}(0,\infty)$ are independent random walks in d-space ($d \geqq 4$), the first starting at O, the second at P where $|P| = \varrho$ and $\frac{1}{2}\varrho^2 < n < 2\varrho^2$; then there are constants c_{21} and $c_{22}^{(d)}$ ($d = 5, 6, \ldots$) such that*

(i) $\mathbf{P}\{\Pi_1^{(4)}(0,n)$ *intersects* $\Pi_2^{(4)}(0,\infty)\} < \dfrac{c_{21}}{\log n}$,

(ii) $\mathbf{P}\{\Pi_1^{(d)}(0,\infty)$ *intersects* $\Pi_2^{(d)}(0,\infty)\} < \dfrac{c_{22}^{(d)}}{n^{(d-4)/2}}$ *for* $d = 5, 6, \ldots$:

PROOF OF (i). Calculate the expected number of points common to $\Pi_1^{(4)}(0,n)$ and $\Pi_2^{(4)}(0,\infty)$. This turns out to be finite. However, if there is an intersection between $\Pi_1^{(4)}\left(0,\left[\frac{n}{2}\right]\right)$ and $\Pi_2^{(4)}(0,\infty)$, both paths can continue independently after the intersection, both of them for at least $\left[\frac{n}{2}\right]$ steps. By Lemma 8 (ii), the conditional expectation of the number of points common is at least $c_{18}\log\frac{n}{2}$ given that an intersection occurs. (i) now follows as otherwise the expected number of common points would not remain finite.

PROOF OF (ii). In the notation of Lemma 7, let $p_1(L), p_2(L)$ be the probabilities that $\Pi_1^{(d)}(0,\infty)$, $\Pi_2^{(d)}(0,\infty)$, respectively, pass through the lattice point L. By using Lemma 1, the result (ii) follows immediately on summing $p_1(L)p_2(L)$ over all the lattice points with integer co-ordinates.

The main result of this section is contained in

THEOREM 1. *Suppose $f(n)$ is an integer-valued function of n which increases to infinity as $n \to \infty$ and $E_n^{(d)}$ is the event that the random walk path in d-space is such that $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+f(n), \infty)$ have at least one point in common.*

(i) *For $d=3$, if $f(n)=n\{\varphi(n)\}^2$ and $\varphi(n)$ is monotonic increasing, then*

$$\mathbf{P}\{E_n^{(3)} \text{ i. o.}\}=0 \quad or \quad 1,$$

according as $\sum_{k=1}^{\infty}\frac{1}{\varphi(2^k)}$ converges or diverges.

(ii) *For $d=4$, if $f(n)=n\psi(n)$ and $\psi(n)$ is monotonic increasing, then*

$$\mathbf{P}\{E_n^{(4)} \text{ i. o.}\}=0 \quad or \quad 1,$$

according as $\sum_{k=1}^{\infty}\frac{1}{k\psi(2^k)}$ converges or diverges.

(iii) *For $d \geqq 5$, if $\sup_{m\geqq n}\frac{f(m)}{m} \leqq c_{23}\frac{f(n)}{n}$,*[1] *then*

$$\mathbf{P}\{E_n^{(d)} \text{ i. o.}\}=0 \quad or \quad 1,$$

according as $\sum_{n=1}^{\infty}\frac{1}{f(n)^{(d-2)/2}}$ converges or diverges.

PROOF OF (i). Our first object will be to obtain

$$\mathbf{P}(E_n^{(3)}) < \frac{c_{24}}{\varphi(n)}, \tag{3.8}$$

for a suitable $c_{24}>0$. Let $Q^{(3)}(n)$ be the number of integers r $(0 \leqq r \leqq 2n)$ such that $\Pi^{(3)}(n+f(n), \infty)$ returns to the point $S_3(r)$. By (2.2) it is clear that

$$\mathscr{E}\{Q^{(3)}(n)\} \leqq c_{24}\sum_{r=0}^{2n}\frac{1}{[n-r+f(n)]^{1/2}}.$$

Thus

$$\mathscr{E}\{Q^{(3)}(n)\} < \frac{c_{25}\,n^{1/2}}{\varphi(n)}, \tag{3.9}$$

since $f(n)=n\{\varphi(n)\}^2$ and $\varphi(n) \to +\infty$ as $n \to \infty$.

As in the proof of Lemma 9 (i) we can estimate $\mathscr{E}\{Q^{(3)}(n)\}$ by another method. If $E_n^{(3)}$ occurs, this means that there exist integers r_1, r_2 with $0 \leqq r_1 \leqq n$ and $r_2 > n+f(n)$ such that $S_3(r_1)=S_3(r_2)$. Now think of $\Pi^{(3)}(r_1, 2n)$ and

[1] This condition is not really necessary for the truth of the theorem. It is inserted because it simplifies the proof slightly.

$\Pi^{(3)}(r_2, \infty)$ as two independent random walks of length $\geqq n$ starting from the same point. Since $r_2 > n\{\varphi(n)\}^2$ and $\varphi(n) \to \infty$, the knowledge that $S_3(r_2) = S_3(r_1)$ will have no appreciable effect on the behaviour of $\Pi^{(3)}(r_1, 2n)$.[2] Hence by Lemma 8 (i) the conditional expectation of $Q^{(3)}(n)$ given $E_n^{(3)}$ satisfies

$$\mathcal{E}\{Q^{(3)}(n)/E_n^{(3)}\} > c_{26} n^{1/2}.$$

Hence

$$\mathcal{E}\{Q^{(3)}(n)\} \geqq \mathbf{P}(E_n^{(3)})\mathcal{E}\{Q^{(3)}(n)/E_n^{(3)}\} > c_{26} n^{1/2}\mathbf{P}(E_n^{(3)}).$$

Using (3.9), this immediately establishes (3.8).

For $k = 2, 3, \ldots$, let $R_k^{(3)}$ be the event that $\Pi^{(3)}(0, 2^k)$ and $\Pi^{(3)}(2^k + 2^{k-1}\{\varphi(2^{k-1})\}^2, \infty)$ have a point in common. By (3.8)

$$\mathbf{P}(R_k^{(3)}) < \frac{2c_{24}}{\varphi(2^{k-1})} \qquad (k = 2, 3, \ldots). \tag{3.10}$$

Hence, if $\sum \frac{1}{\varphi(2^k)}$ converges, then $\sum \mathbf{P}(R_k^{(3)})$ converges and, by Borel—Cantelli,

$$\mathbf{P}\{R_k^{(3)} \ i.o.\} = 0.$$

But from the definition of $R_k^{(3)}$, the non-occurrence of $R_k^{(3)}$ implies that no $E_n^{(3)}$ occurs for $n_{k-1} \leqq n \leqq n_k$. Hence

$$\mathbf{P}\{E_n^{(3)} \ i.o.\} = 0.$$

Conversely, suppose $\sum \frac{1}{\varphi(2^k)}$ diverges. We have independence difficulties in applying Borel—Cantelli this time: instead we prove that there exists $\eta > 0$ such that

$$\mathbf{P}\{E_n^{(3)} \ i.o.\} > \eta. \tag{3.11}$$

By the law of 0 or 1 this implies $\mathbf{P}\{E_n^{(3)} \ i.o.\} = 1$, and so completes the proof of (i).

Let $F_k^{(3)}$ $(k = 2, 3, \ldots)$ be the event that $\Pi^{(3)}\left(\frac{1}{2}n, n\right)$ and $\Pi^{(3)}(n + f(n), n + 5f(n))$ have a point in common for $n = 2^k$. Let $H_k^{(3)}$ $(k = 2, 3, \ldots)$ be the event that $\Pi^{(3)}(n + f(n), n + 4f(n))$ returns within $\left(\frac{1}{2}n\right)^{1/2}$ of $S_3\left(\frac{1}{2}n\right)$

[2] Strictly speaking the conditional probability distribution for $\{S_3(r) - S_3(r_1)\}$ $(r_1 \leqq r \leqq 2n)$ given $S_3(r_2) = S_3(r_1)$, should be obtained. It can be shown by simple but laborious computation that the conditional probability differs from the 'free' probability by a factor which $\to 1$ as $n \to \infty$; hence it is clearly justifiable to use Lemma 8 (i) with suitable different values for the constants.

for $n=2^k$. By Lemma 4,

$$\mathbf{P}(H_k^{(3)}) > \frac{\frac{1}{10}e_3}{\varphi(2^k)}, \tag{3.12}$$

and after having returned within $\left(\frac{1}{2}n\right)^{1/2}$ of $S_3\left(\frac{1}{2}n\right)$ there are still at least $f(n)>n$ steps to complete $\Pi^{(3)}(n+f(n), n+5f(n))$. Again, since $\varphi(n)\to\infty$, the departure from the situation of two independent random walks each of length $\geqq n$ starting at a distance apart $\sqrt{n}$ becomes negligable for large n. Hence by Lemma 6 (i),

$$\mathbf{P}(F_k^{(3)}/H_k^{(3)}) > c_{27},$$

so that, by (3.12),

$$\mathbf{P}(F_k^{(3)}) > \frac{c_{28}}{\varphi(2^k)}. \tag{3.13}$$

Hence $\sum_{k=2}^{\infty}\mathbf{P}(F_k^{(3)})$ diverges. Since $F_k^{(3)}\subset R_k^{(3)}$, it follows from (3.10) and the fact that $\varphi(n)\to\infty$ that $\mathbf{P}(F_k^{(3)})\to 0$ as $k\to\infty$. Hence for any $\eta>0$ there is a $K=K(\eta)$ such that given $k_1\geqq K$ one can find a $k_2\geqq k_1$ such that

$$2\eta < \sum_{k_1\leqq 4r\leqq k_2}\mathbf{P}(F_{4r}^{(3)}) < 3\eta. \tag{3.14}$$

We shall show that, provided η is chosen small enough, (3.14) will imply

$$\mathbf{P}\{\bigcup_{k_1\leqq k\leqq k_2} F_k^{(3)}\} > \eta \tag{3.15}$$

which clearly implies (3.11).

Suppose now that $F_k^{(3)}$ has happened. This implies that there are integers r_3, r_4 such that $2^{k-1}\leqq r_3\leqq 2^k$, $2^k+2^k\{\varphi(2^k)\}^2\leqq r_4\leqq 2^k+5\cdot 2^k\{\varphi(2^k)\}^2$ and $S_3(r_3)=S_3(r_4)$. In considering whether or not $F_{k+4}^{(3)}$ happens, there are at least

$$(16-5)2^k\{\varphi(2^{k+4})\}^2 > \frac{1}{2}f(2^{k+4})$$

steps after r_4 before we look at $\Pi^{(3)}(n+f(n), n+5f(n))$ for $n=2^{k+4}$. The method used earlier for estimating $R_k^{(3)}$ shows immediately that

$$\mathbf{P}\{F_{k+4}^{(3)}/F_k\} < \frac{c_{29}}{\varphi(2^{k+4})},$$

so that, by (3.13), it follows that

$$\mathbf{P}\{F_{k+4}^{(3)}\cap F_k^{(3)}\} < c_{30}\mathbf{P}(F_k)\mathbf{P}(F_{k+4}).$$

The same argument shows that for an integer $l \geqq 4$,

$$\mathbf{P}(F_{k+l} \cap F_k) < c_{30} \mathbf{P}(F_k) \mathbf{P}(F_{k+l}). \tag{3.16}$$

If (3. 14) is satisfied, (3. 16) implies that

$$\mathbf{P}\{F_{4r} \cap \bigcup_{4r < 4t \leqq k_2} F_{4t}\} < 3\eta c_{30} \mathbf{P}(F_{4r}),$$

so that if η is chosen small enough to ensure that $3\eta c_{30} < \frac{1}{2}$, it follows that for $k_1 \leqq 4r \leqq k_2$

$$\mathbf{P}\{F_{4r} - F_{4r} \cap \bigcup_{4r < 4t \leqq k_2} F_{4t}\} > \frac{1}{2} \mathbf{P}(F_{4r}).$$

Using (3. 14), this clearly establishes (3. 15) as required.

Proof of (ii). A very similar argument will work. Instead of (3. 10), (3. 13) one obtains

$$\mathbf{P}(R_k^{(4)}) < \frac{c_{31}}{(k-1)\psi(2^{k-1})}, \quad \mathbf{P}(F_k^{(4)}) > \frac{c_{32}}{k\varphi(2^k)}.$$

Proof of (iii). Define a sequence of integers by: n_1 is some fixed integer such that $f(n_1) \geqq 4$,

$$n_{k+1} = n_k + \left[\frac{1}{2} f(n_k)\right] \quad (k = 1, 2, \ldots). \tag{3.17}$$

Under the conditions given it is clear that $\sum_{n=1}^{\infty} \frac{1}{f(n)^{(d-2)/2}}$ converges or diverges with the series

$$\sum_{k=1}^{\infty} \frac{1}{f(n_k)^{(d-4)/2}}. \tag{3.18}$$

Suppose first that (3. 18) converges. For $d \geqq 5$, let $Q^{(d)}(k)$ be the number of points common to $\Pi^{(d)}(0, n_k)$ and $\Pi^{(d)}\left(n_k + \left[\frac{1}{2} f(n_{k-1})\right], \infty\right)$. Then the argument used in (i) shows that

$$\mathcal{E}\{Q^{(d)}(k)\} < \frac{c_{33}}{\{f(n_{k-1})\}^{(d-4)/2}}.$$

Hence

$$\mathbf{P}\{Q^{(d)}(k) \geqq 1\} < \frac{c_{33}}{\{f(n_{k-1})\}^{(d-4)/2}}.$$

By Borel—Cantelli, there is probability 1 that $Q^{(d)}(k) = 0$ except for finitely

many values of k. Further, if $Q^{(d)}(k)=0$, then $E_n^{(d)}$ does not happen for $n_{k-1}\leqq n\leqq n_k$. Thus

$$\mathbf{P}\{E_n^{(d)}\ i.\,o.\}=0.$$

Conversely suppose that (3. 18) diverges. Let $F_k^{(d)}$ be the event that there is a point common to $\Pi^{(d)}(n_k-f(n_k), n_k)$ and $\Pi^{(d)}(n_k+f(n_k), n_k+ +2f(n_k))$. By (3. 17), under the conditions satisfied by $f(n)$, the events $F_k^{(d)}$ and $F_{k+l}^{(d)}$ are completely independent when $l\geqq c_{34}$ where c_{34} is a suitable integer. Hence it is sufficient to prove that $\sum_{r=1}^{\infty}\mathbf{P}(F_{rc_{34}}^{(d)})$ diverges and apply Borel—Cantelli.

Now if $H_k^{(d)}$ is the event

$$|S_d(n_k)-S_d(n_k+f(n_k))|<(f(n_k))^{1/2},$$

then an easy computation shows that $\mathbf{P}(H_k^{(d)})>c_{35}^{(d)}$. Further, by Lemma 6 (iii),

$$\mathbf{P}(F_k^{(d)}/H_k^{(d)})>\frac{c_6^{(d)}}{\{f(n_k)\}^{(d-4)/2}},$$

so that

$$\mathbf{P}(F_k^{(d)})>\frac{c_{36}^{(d)}}{\{f(n_k)\}^{(d-4)/2}}.$$

Since $f(n)$ is increasing with n and (3. 18) diverges, it follows that $\sum_{r=1}^{\infty}\mathbf{P}(F_{rc_{34}}^{(d)})$ also diverges. This completes the proof of the theorem.

Remark. The result of Theorem 1 (ii) implies that there are infinitely many values of n such that $\Pi^4(0, n)$ and $\Pi^4(2n, \infty)$ have points in common. A somewhat simplified version of the same proof is sufficient to prove the following interesting result:

Theorem 2. (i) *Two independent infinite random walks in 4-space which start from any two given fixed points have infinitely many common points with probability* 1; *whereas two random walk paths in d-space* ($d\geqq 5$) *meet only finitely often, with probability* 1.

(ii) *Three independent infinite random walks in 3-space which start from any 3 given fixed points have infinitely many points in common with probability* 1; *whereas three random walk paths in d-space* ($d\geqq 4$) *have only finitely many common points, with probability* 1.

The relationship between these results and those for Brownian motion was discussed in the introduction.

4. Points where the past and future do not intersect. We first need to estimate the probability that two random walk paths starting from the origin have no points (other than their starting point) in common.

For $d \geqq 3$, let $\tau_d(n)$ be the probability that two independent paths of n steps in d-space starting from O have no intersection. Let $\mu_d(n)$ be the corresponding probability for two paths, one of n steps, the other of infinite length. Clearly,

$$\begin{aligned} &\tau_d(1) \geqq \tau_d(2) \geqq \cdots \geqq \tau_d(n) \geqq \dots, \\ &\mu_d(1) \geqq \mu_d(2) \geqq \cdots \geqq \mu_d(n) \geqq \dots \end{aligned} \tag{4.1}$$

and

$$\tau_d(n) \geqq \mu_d(n) \qquad (n = 1, 2, \dots). \tag{4.2}$$

Let $S_d(1), S_d(2), \dots, S_d(n)$ be the sequence of points in path Π_1, and $S_d^1(1), S_d^1(2), \dots$ the points of the path Π_2. Put

$$\varrho_d(r) = |S_d(r)|.$$

Then we can enumerate the possibilities that $S_d(r)$ is entered by Π_2, but later points of Π_1 are not entered. This gives

$$\tau_d(n) \sum_{r=0}^{n} p_r \leqq 1 \tag{4.3}$$

where p_r is the probability that Π_2 enters $S_d(r)$ before the n^{th} step, since the probability of later non-intersection is greater than that for the non-intersection of two paths of n-steps.

For $d = 3$, by Lemma 2,

$$p_r > c_{37} \mathfrak{E} \left\{ \frac{1}{1 + \varrho_3(r)} \right\} > \frac{c_{38}}{r^{1/2}}, \quad \text{by Lemma 6.}$$

Hence by (4.3) we have

$$\tau_3(n) \leqq \frac{c_{39}}{n^{1/2}}. \tag{4.4}$$

For $d = 4$, by Lemma 2,

$$p_r > c_{40} \mathfrak{E} \left\{ \frac{1}{1 + \{\varrho_4(r)\}^2} \right\} > \frac{c_{41}}{r}, \quad \text{by Lemma 5.}$$

Hence by (4.3) we have

$$\tau_4(n) \leqq \frac{c_{42}}{\log n}. \tag{4.5}$$

To obtain estimates in the other direction, it is easier to use $\mu_d(n)$. Clearly, for $0 \leqq k_1 \leqq n$,

$$\mu_d(n-k_1)\sum_{k=1}^{k_1} q_k + \sum_{k=k_1+1}^{n} q_k \geqq 1, \tag{4.6}$$

where q_k is the probability that Π_2 enters $S_d(k)$ at least once.

For $d=3$, the method does not give a good lower bound for $\mu_3(n)$. More complicated computations show that $\mu_3(n) \geqq c_{43} n^{1/2}$, but we do not prove this as it is not required in the sequel.

For $d=4$, take $k_1 = n - \left[\dfrac{n}{\log n}\right]$. Then, if $S_4(k)=0$, Π_2 certainly enters $S_4(k)$ at least once. If $S_4(k) \neq 0$, then by Lemma 1, the probability that Π_2 enters $S_4(k)$ is at most $c_{44}/\{\varrho_4(k)\}^2$. Since $\mathbf{P}\{S_4(k)=0\} < c_{45}/k^2$, it follows that $q_k < \dfrac{c_{45}}{k^2} + 2c_{44}\mathscr{E}\left\{\dfrac{1}{1+\{\varrho_4(k)\}^2}\right\}$. Using Lemma 5 this gives $q_k < \dfrac{c_{46}}{k}$. Substituting this in (4.6) gives immediately

$$\mu_4\left(\left[\frac{n}{\log n}\right]\right) c_{46} \sum_{k=1}^{n} \frac{1}{k} + O\left(\frac{1}{\log n}\right) \geqq 1$$

which implies immediately

$$\mu_4(n) \geqq \frac{c_{47}}{\log n}. \tag{4.7}$$

For $d \geqq 5$, a similar argument shows that $\sum_{k=1}^{\infty} q_k$ is a convergent series. Hence, if $k_1 = \left[\dfrac{1}{2} n\right]$,

$$\mu_d\left(\left[\frac{1}{2} n\right]\right) c_{48} + o(1) \geqq 1$$

which implies that

$$\mu_d(n) \geqq c_{49}^{(d)}. \tag{4.8}$$

By (4.1), (4.2) it follows that for $d \geqq 5$, there are constants $\tau_d \geqq \mu_d > 0$ such that

$$\tau_d(n) = \tau_d(1+o(1)). \tag{4.9}$$

In fact, the results of Section 3 imply that $\tau_d = \mu_d$ for $d \geqq 5$, and further that the probability of the non-intersection of two infinite random walks starting from the origin is also τ_d.

THEOREM 3. *For* $d \geqq 3$, *let* $G^{(d)}(n)$ *be the number of integers* r $(1 \leqq r \leqq n)$ *for which* $\Pi^{(d)}(0, r)$ *and* $\Pi^d(r+1, \infty)$ *have no points in common, then*

(i) *for any* $\varepsilon > 0$,

$$\mathbf{P}\{G^{(3)}(n) > n^{1/2+\varepsilon} \quad i.\,o.\} = 0;$$

(ii) $$\mathbf{P}\left\{0 = \liminf_{n\to\infty} \frac{G^{(4)}(n)\log n}{n} \leqq \limsup_{n\to\infty} \frac{G^{(4)}(n)\log n}{n} \leqq c_{50}\right\} = 1;$$

(iii) *for* $d \geqq 5$,

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{G^{(d)}(n)}{n} = \tau_d\right\} = 1.$$

PROOF OF (i). By (4.3), we have $\mu_3(n) \leqq c_{39} n^{-1/2}$, so that

$$\mathcal{E}\{G^{(3)}(n)\} \leqq c_{39} \sum_{r=1}^{n} r^{-1/2} < c_{51} n^{1/2}.$$

Hence

$$\mathbf{P}\left\{G^{(3)}(n) > \frac{1}{2} n^{1/2+\varepsilon}\right\} < \frac{c_{52}}{n^{\varepsilon}},$$

so that if $n_k = [k^{2/\varepsilon}]$, the series

$$\sum \mathbf{P}\left\{G^{(3)}(n_k) > \frac{1}{2} n_k^{1/2+\varepsilon}\right\}$$

converges and there is probability 1 that there exists k_0 such that $G^{(3)}(n_k) \leqq \frac{1}{2} n_k^{1/2+\varepsilon}$ for $k \geqq k_0$. Since $G^{(3)}(n)$ increases with n, and $\frac{n_{k+1}}{n_k} \to 1$ as $k \to \infty$, this implies that for $n_k \leqq n \leqq n_{k+1}$, $k \geqq k_0$, $G^{(3)}(n) \leqq n^{1/2+\varepsilon}$.

PROOF OF (ii). Let $T^{(4)}(r)$ be the event that $\Pi^{(4)}\left(r - \left[\frac{r}{(\log r)^4}\right], r\right)$ and $\Pi^{(4)}\left(r+1, r+\left[\frac{r}{(\log r)^4}\right]\right)$ have no point in common and let $H^{(4)}(n)$ be the number of events $T^{(4)}(r)$ which occur for $1 \leqq r \leqq n$. Clearly,

$$H^{(4)}(n) \geqq G^{(4)}(n). \tag{4.10}$$

However, by (4.5) and (4.7),

$$\frac{c_{53}}{\log r} \leqq \mathbf{P}\{T^{(4)}(r)\} \leqq \frac{c_{54}}{\log r},$$

so that

$$c_{53} \frac{n}{\log n} \leqq \mathcal{E}\{H^{(4)}(n)\} \leqq c_{55} \frac{n}{\log n}. \tag{4.11}$$

Now the events $T^{(4)}(k)$ and $T^{(4)}(r)$ are completely independent whenever

$k>r+4\frac{n}{(\log n)^4}$, so that the variance $\sigma^2\{H^{(4)}(r)\}$ of $H^{(4)}(n)$ satisfies

$$\sigma^2\{H^{(4)}(n)\} \leqq c_{56}\frac{n^2}{(\log n)^4}. \tag{4.12}$$

By Chebyshev's inequality, using (4.11) and (4.12), we have

$$\mathbf{P}\left\{|H^{(4)}(n)-\mathcal{E}\{H^{(4)}(n)\}|>\frac{\varepsilon n}{\log n}\right\}<\frac{c_{57}}{\varepsilon^2}\frac{1}{(\log n)^2}. \tag{4.13}$$

For the sequence $m_k=[e^{k/\log k}]$ an application of the Borel—Cantelli lemma shows that for every $\varepsilon>0$

$$|H^{(4)}(m_k)-\mathcal{E}\{H^{(4)}(m_k)\}|\leqq\frac{\varepsilon m_k}{\log m_k}$$

except for finitely many integers k with probability 1. Since $\frac{m_{k+1}}{m_k}\to 1$ as $k\to\infty$, this implies that

$$\lim_{n\to\infty}\frac{\log n}{n}|H^{(4)}(n)-\mathcal{E}\{H^{(4)}(n)\}|=0$$

with probability 1. By (4.11), (4.10) it follows immediately that

$$\mathbf{P}\left\{\limsup_{n\to\infty}\frac{\log n}{n}H^{(4)}(n)\leqq c_{55}\right\}=1,$$

which by (4.10) implies the right side inequality of (ii).

By Theorem 1 (ii), there are, with probability 1, infinitely many integers n for which

$$\Pi^{(4)}(0, n) \quad\text{and}\quad \Pi^{(4)}(n+n[\log\log n], \infty)$$

have a point in common. For such values of n there are no points of non-intersection between n and $n+n[\log\log n]$, so that $G^{(4)}(n)=G^{(4)}(n[\log\log n])$. Thus it follows immediately from the right side inequality of (ii) that

$$\mathbf{P}\left\{\liminf_{n\to\infty}\frac{\log n}{n}G^{(4)}(n)=0\right\}=1.$$

PROOF OF (iii). For $d\geqq 5$, let $T^{(d)}(r)$ be the event that $\Pi^{(d)}(r-[r^{3/4}], r)$ and $\Pi^{(d)}(r+1, r+[r^{3/4}])$ have no points in common and let $H^{(d)}(r)$ be the number of events $T^{(d)}(r)$ $(1\leqq r\leqq n)$ which occur. By (4.9),

$$\mathbf{P}\{T^{(d)}(r)\}=\tau_d(1+o(1)),$$

so that

$$\mathcal{E}\{H^{(d)}(r)\}=n\tau_d(1+o(1)). \tag{4.14}$$

If $k>r+4n^{3/4}$, the events $T^{(d)}(r)$ and $T^{(d)}(k)$ are completely independent. Hence

$$\sigma^2\{H^{(d)}(n)\}<c_{58}n^{7/4}. \tag{4.15}$$

The usual Chebyshev estimate using the sequence $r_k = k^{10}$ $(k = 1, 2, \ldots)$ now establishes using (4. 14), (4. 15) that

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{H^{(d)}(n)}{n} = \tau_d\right\} = 1. \tag{4.16}$$

However, from Theorem 1 (iii) there is, with probability 1, an integer n_0 such that for $n \geqq n_0$ (a) $\Pi^{(d)}(0, n-[n^{3/4}])$ and $\Pi^{(d)}(n, \infty)$ do not intersect and (b) $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+[n^{3/4}], \infty)$ do not intersect. For $n \geqq n_0$ the difference $G^{(d)}(n) - H^{(d)}(n)$ remains fixed, so that

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{G^{(d)}(n) - H^{(d)}(n)}{n} = 0\right\} = 1.$$

This result, together with (4. 16), establishes

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{G^{(d)}(n)}{n} = \tau_d\right\} = 1,$$

and completes the proof of the theorem.

REMARK 1. The result of Theorem 3 (i) is clearly not best-possible: even the methods we have used can give a better upper bound for the density $\frac{1}{n} G^{(3)}(n)$. We expect that $\frac{1}{n^{1/2}} G^{(3)}(n)$ is bounded but did not succeed in proving this.

REMARK 2. An obvious problem arises out of the result of Theorem 2 (ii). Is $\limsup_{n\to\infty} \frac{\log n}{n} G^{(4)}(n)$ positive or zero?

REMARK 3. It is clear that τ_d, the density of good points, increases with d. In fact, $\tau_d \to 1$ as $d \to \infty$.

(Received 16 November 1959)

References

[1] A. DVORETZKY and P. ERDŐS, Some problems on random walk in space, *Proc. Second Berkeley Symposium*, (1950), pp. 353–368.

[2] A. DVORETZKY, P. ERDŐS and S. KAKUTANI, Double points of Brownian paths in n-space, *Acta Sci. Math. Szeged*, **12** (1950), pp. 75–81.

[3] A. DVORETZKY, P. ERDŐS, S. KAKUTANI and S. J. TAYLOR, Triple points of Brownian paths in 3-space, *Proc. Cambr. Phil. Soc.*, **53** (1957), pp. 856–862.

[4] P. ERDŐS and S. J. TAYLOR, Some problems concerning the structure of random walk paths, *Acta Math. Acad. Sci. Hung.*, **11** (1960), pp. 137–162.

[5] S. J. TAYLOR, The Hausdorff α-dimensional measure of Brownian paths in n-space, *Proc. Cambr. Phil. Soc.*, **49** (1953), pp. 31–39.

ON SOME INFINITE SERIES AND A CERTAIN LIMIT PROPERTY OF REAL ENTIRE FUNCTIONS

By

M. KUCZMA (Kraków)

(*Presented by* P. TURÁN)

In my paper [1] I discussed the functions[1]

$$S_p(x) = \sum_{n=1}^{\infty} \frac{n^p}{x^n} \qquad (x > 1;\ p = 0, 1, 2, \ldots)$$

and from properties of the functions $S_p(x)$ I deduced the following theorem on real entire functions:

For every real entire function $f(t)$

$$\limsup_{n\to\infty} \sqrt[n]{|f^{(n)}(0)|} \geqq \ln \limsup_{n\to\infty} \sqrt[n]{|f(n)|}. \tag{1}$$

If, moreover, $f(t) \not\equiv 0$ *and* $f^{(n)}(0) \geqq 0$ $(n = 0, 1, 2, \ldots)$, *then in the relation* (1) *the equality holds.*

Recently Prof. P. TURÁN has raised the problem whether in the right-hand side of the relation (1) the sequence $f(n)$ can be replaced by $f(\alpha_n)$, where α_n is any positive sequence such that

$$\lim_{n\to\infty} \frac{\alpha_n}{n} = 1.$$

The answer turns out to be positive, and the method used to prove the relation (1) is also applicable. However, as the paper [1] is not readily available, I have decided to publish this slight generalization of the original result.

1. Given a sequence α_n with positive terms, fulfilling the relation

$$\lim_{n\to\infty} \frac{\alpha_n}{n} = A \qquad (0 < A < \infty),$$

we shall denote by $T_p(x)$ the sum of the series

$$T_p(x) = \sum_{n=1}^{\infty} \alpha_n^p e^{-nx} \qquad (x > 0;\ p = 0, 1, 2, \ldots).$$

I shall prove the following

[1] The functions $S_p(x)$ were also investigated by P. VERMES ([2], p. 79).

THEOREM I. *For every* $x>0$

$$\lim_{p\to\infty}\sqrt[p]{\frac{T_p(x)}{p!}}=\frac{A}{x}. \tag{2}$$

PROOF. Let us assume at first that $A=1$.
We consider the series

$$\sum_{p=0}^{\infty}\sum_{n=1}^{\infty}\frac{t^p\alpha_n^p}{p!}e^{-nx}=\sum_{p=0}^{\infty}\frac{t^p}{p!}T_p(x) \tag{3}$$

where t is a parameter, $t>0$. Since the terms of the series (3) are positive, we can change the order of summation. We obtain

$$\sum_{n=1}^{\infty}\sum_{p=1}^{\infty}\frac{t^p\alpha_n^p}{p!}e^{-nx}=\sum_{n=1}^{\infty}e^{-(nx-t\alpha_n)}. \tag{4}$$

We put $\delta\overset{\text{def}}{=}x-t$. Thus the series (4) can be written as

$$\sum_{n=1}^{\infty}e^{n\left[t\left(\frac{\alpha_n}{n}-1\right)-\delta\right]}. \tag{5}$$

For n sufficiently large we have

$$\left|\frac{\alpha_n}{n}-1\right|<\frac{|\delta|}{2t}.$$

Hence, if $\delta>0$,

$$\frac{\alpha_n}{n}-1<\frac{\delta}{2t},$$

whence

$$t\left(\frac{\alpha_n}{n}-1\right)-\delta<-\frac{\delta}{2},$$

and consequently the series (5) converges. On the other hand, if $\delta<0$, then we have by the inequality $1-\frac{\alpha_n}{n}\leq\left|\frac{\alpha_n}{n}-1\right|$:

$$1-\frac{\alpha_n}{n}<-\frac{\delta}{2t},$$

whence

$$t\left(\frac{\alpha_n}{n}-1\right)-\delta>-\frac{\delta}{2},$$

and consequently the series (5) diverges. Thus the series (5) converges for $x>t$ and diverges for $x<t$. Consequently, also the series (3) converges for

$x>t$ and diverges for $x<t$. Hence we obtain the inequalities

$$t \limsup_{p\to\infty} \sqrt[p]{\frac{T_p(x)}{p!}} \leqq 1 \quad \text{for} \quad x>t$$

and

$$t \limsup_{p\to\infty} \sqrt[p]{\frac{T_p(x)}{p!}} \geqq 1 \quad \text{for} \quad x<t$$

whence

$$\limsup_{p\to\infty} \sqrt[p]{\frac{T_p(x)}{p!}} = \frac{1}{x}. \tag{6}$$

Now, from the evident inequality

$$T_p(x) \geqq \alpha^p_{\left[\frac{p}{x}\right]} e^{-\left[\frac{p}{x}\right]x}$$

($[a]$ denotes the greatest integer not exceeding a) we have

$$\sqrt[p]{\frac{T_p(x)}{p!}} \geqq \frac{1}{p}\alpha_{\left[\frac{p}{x}\right]} \sqrt[p]{\frac{p^p}{p!}} e^{-\frac{1}{p}\left[\frac{p}{x}\right]x} = \frac{1}{p}\left[\frac{p}{x}\right] \sqrt[p]{\frac{p^p}{p!}} \frac{\alpha_{\left[\frac{p}{x}\right]}}{\left[\frac{p}{x}\right]} e^{-\frac{1}{p}\left[\frac{p}{x}\right]x}.$$

Hence, on account of the relation

$$\lim_{p\to\infty} \frac{1}{p}\left[\frac{p}{x}\right] = \frac{1}{x},$$

we obtain

$$\liminf_{p\to\infty} \sqrt[p]{\frac{T_p(x)}{p!}} \geqq \frac{1}{x}. \tag{7}$$

From (6) and (7) it follows

$$\lim_{p\to\infty} \sqrt[p]{\frac{T_p(x)}{p!}} = \frac{1}{x}.$$

Now, let A be arbitrary, $0<A<\infty$. We put

$$\beta_n = \frac{\alpha_n}{A},$$

and

$$R_p(x) = \sum_{n=1}^{\infty} \beta_n^p e^{-nx} \qquad (x>0;\ p=0, 1, 2, \ldots).$$

On account of what has just been proved

$$\lim_{p\to\infty}\sqrt[p]{\frac{R_p(x)}{p!}}=\frac{1}{x}$$

whence the relation (2) follows. This completes the proof.

Now, let u_p $(p=0,1,2,\ldots)$ be an arbitrary sequence of real numbers. A consequence of the above theorem is the following

THEOREM II. *The series*

$$\sum_{p=0}^{\infty}u_p T_p(x)$$

converges absolutely for

$$x>A\limsup_{p\to\infty}\sqrt[p]{|u_p|p!}\ \textit{and diverges for}\ \ x<A\limsup_{p\to\infty}\sqrt[p]{|u_p|p!}.$$

PROOF. We have

$$\limsup_{p\to\infty}\sqrt[p]{|u_p T_p(x)|}=\lim_{p\to\infty}\sqrt[p]{\frac{T_p(x)}{p!}}\limsup_{p\to\infty}\sqrt[p]{|u_p|p!}=\frac{A}{x}\limsup_{p\to\infty}\sqrt[p]{|u_p|p!}$$

whence

$$\limsup_{p\to\infty}\sqrt[p]{|u_p T_p(x)|}<1\quad\text{for}\quad x>A\limsup_{p\to\infty}\sqrt[p]{|u_p|p!}$$

and

$$\limsup_{p\to\infty}\sqrt[p]{|u_p T_p(x)|}>1\quad\text{for}\quad x<A\limsup_{p\to\infty}\sqrt[p]{|u_p|p!}.$$

This completes the proof.

2. Now I shall prove the following

THEOREM III. *Let $f(t)$ be an arbitrary real entire function. Then*

$$A\limsup_{n\to\infty}\sqrt[n]{|f^{(n)}(0)|}\geqq\ln\limsup_{n\to\infty}\sqrt[n]{|f(\alpha_n)|}. \tag{8}$$

If, moreover, $f(t)\not\equiv 0$ and $f^{(n)}(0)\geqq 0$ $(n=0,1,2,\ldots)$, then in the relation (8) *the equality holds.*

PROOF. We form the series

$$\sum_{p=0}^{\infty}u_p T_p(x)=\sum_{p=0}^{\infty}\sum_{n=1}^{\infty}u_p\alpha_n^p e^{-nx} \tag{9}$$

where $u_p=\frac{1}{p!}f^{(p)}(0)$. This series converges absolutely for

$$x>A\limsup_{p\to\infty}\sqrt[p]{|u_p|p!}=A\limsup_{p\to\infty}\sqrt[p]{|f^{(p)}(0)|}. \tag{10}$$

Thus for x fulfilling (10) we can change the order of summation in the double series (9). Consequently, the series

$$\sum_{n=1}^{\infty}\sum_{p=0}^{\infty} u_p \alpha_n^p e^{-nx} = \sum_{n=1}^{\infty} f(\alpha_n) e^{-nx} \tag{11}$$

converges for x fulfilling (10). But, on the other hand, the series (11) is known to converge for

$$x > \ln \limsup_{n\to\infty} \sqrt[n]{|f(\alpha_n)|}. \tag{12}$$

Hence the inequality (8) follows.

If $u_p \geqq 0$ $(p = 0, 1, 2, \ldots)$, then the double series (9) and (11), as series with positive terms, simultaneously converge or diverge. Thus both the right-hand sides of (10) and (12) must be equal. This completes the proof.

(Received 16 November 1959)

Bibliography

[1] M. Kuczma, O sumach szeregów $\sum_{n=1}^{\infty} \frac{n^p}{x^n}$, *Zeszyty Naukowe Uniw. Jagiell., Mat-Fiz-Chem.*, 4 (1958), pp. 39–46.

[2] P. Vermes, The transpose of a summability matrix, Colloque sur la théorie des suites, tenu à Bruxelles du 18 au 20 décembre 1957 (Paris—Louvain, 1958), pp. 60–86.

О ТРИГОНОМЕТРИЧЕСКОМ (0, 2)-ИНТЕРПОЛИРОВАНИИ

О. КИШ (Будапешт)

(Представлено П. Тураном)

§ 1

Настоящая заметка примыкает к ряду работ, посвященных теории (0, 2)-интерполяционных многочленов. Так инициатор изучения этих многочленов П. Туран назвал многочлены $R_n(x)$ не выше $2n-1$-ой степени, удовлетворяющие в узлах $x_1, x_2, \ldots, x_n$ условиям

$$\left.\begin{aligned} R_n(x_k) &= \alpha_k \\ R_n''(x_k) &= \beta_k \end{aligned}\right\} \qquad (k=1, 2, \ldots, n).$$

Здесь α_k и β_k некоторые заданные числа.

В связи с (0, 2)-интерполяционными многочленами были поставлены следующие задачи:

1. Для данных узлов интерполирования существует ли единственный (0, 2)-интерполяционный многочлен при любых значениях чисел α_k и β_k?

2. Пусть для некоторых узлов предыдущая задача решается положительно. Найти явный вид фундаментальных многочленов интерполирования, т. е. таких многочленов не выше $2n-1$-ой степени, которые удовлетворяют условиям

$$r_k(x_j) = \begin{cases} 1, & \text{если} \quad j=k, \\ 0, & \text{если} \quad j\neq k, \end{cases}$$

$$\begin{aligned} r_k''(x_j) &= 0, \\ \varrho_k(x_j) &= 0, \end{aligned} \qquad (j, k=1, 2, \ldots, n),$$

$$\varrho_k''(x_j) = \begin{cases} 1, & \text{если} \quad j=k, \\ 0, & \text{если} \quad j\neq k. \end{cases}$$

(С их помощью (0, 2)-интерполяционный многочлен можно представить формулой

$$R_n(x) = \sum_{k=1}^{n} [\alpha_k r_k(x) + \beta_k \varrho_k(x)].)$$

3. Пусть для бесконечной последовательности систем узлов $x_{1,n}$, $x_{2,n}, \ldots, x_{n,n}$ первая задача решается положительно. Обозначим через $r_{k,n}(x)$ и $\varrho_{k,n}(x)$ фундаментальные многочлены, соответствующие этим узлам.

Какие условия следует наложить на функцию $f(x)$ и числа $\beta_{k,n}$ для того, чтобы последовательность многочленов

$$\sum_{k=1}^{n}[f(x_{k,n})r_{k,n}(x)+\beta_{k,n}\varrho_{k,n}(x)] \tag{1}$$

равномерно сходилась бы к $f(x)$?

Приведем полученные до сих пор результаты, относящиеся к этим задачам.

Если вещественные узлы расположены симметрично относительно нуля и число их нечетно, то первая задача решается отрицательно. Пусть узлы суть корни ультрасферического многочлена $P_n^{(\lambda)}(x)$. Если $\lambda+\frac{1}{2}$ четное число и $n\geqq\lambda+\frac{5}{2}$, то эта задача также имеет отрицательное решение, но если $\lambda\geqq-\frac{1}{2}$, $\lambda+\frac{1}{2}$ не есть четное число и $n\geqq 4$ четно, то первая задача решается положительно.[1] Она имеет положительное решение и в том случае, если $n\geqq 2$ и узлами являются числа $\exp i\frac{2\pi}{n}k$.[2]

Явный вид фундаментальных многочленов получен для случая, когда узлы суть корни многочлена

$$\Pi_n(x)=(1-x^2)P'_{n-1}(x),$$

где $P_{n-1}(x)$ многочлен Лежандра $n-1$-ой степени и их число четно,[3] а также для случая узлов $\exp i\frac{2\pi}{n}k$.[2]

Для этих же узлов исследована сходимость интерполяционного процесса. В первом случае многочлены (1) сходятся равномерно к $f(x)$ на отрезке $[-1, +1]$, если функция $f(x)$ удовлетворяет условию

$$f(x+h)-2f(x)+f(x-h)=o(h) \qquad (-1\leqq x-h<x+h\leqq+1), \tag{2}$$

а числа $\beta_{k,n}$ условию

$$\beta_{1,n}=o(n^2), \quad \beta_{k,n}=\frac{o(n)}{\sqrt{1-x_{k,n}^2}} \quad (k=2,3,\ldots,n-1), \quad \beta_{n,n}=o(n^2).$$

[1] J. Surányi and P. Turán, Notes on interpolation. I, *Acta Math. Acad. Sci. Hung.*, **6** (1955), стр. 67—79.

[2] О. Киш, Замечания об интерполировании, *Acta Math. Acad. Sci. Hung.*, **11** (1960), стр. 49—64.

[3] J. Balázs and P. Turán, Notes on interpolation. II, *Acta Math. Acad. Sci. Hung.*, **8** (1957), стр. 201—215.

Условие (2) не может быть заменено условием Липшица с показателем $\alpha<1$, даже если все $\beta_{k,n}$ равны нулю.[4,5]

Во втором случае для сходимости многочленов (1) в круге $|x|\leqq 1$ достаточно непрерывности функции $f(x)$ в этом круге, ее аналитичности при $|x|<1$, выполнения условия Дини—Липшица для модуля непрерывности **функции $f(\exp ix)$ (здесь x вещественно)** и соотношения

$$\beta_{k,n}=o\left(\frac{n^2}{\log n}\right) \qquad (k=1, 2, \ldots, n).$$

В настоящей заметке задачи, о которых говорилось выше, решаются для случая, когда интерполирование производится тригонометрическими многочленами не выше чем n-ого порядка, а в качестве узлов интерполирования выбираются точки

$$x_k=\frac{2\pi}{n}k \qquad (k=0, 1, \ldots, n-1).$$

Доказываются следующие теоремы:

Теорема 1. Если число узлов четно, то (0, 2)-интерполяционный многочлен, вообще говоря, не существует.

Теорема 2. Если число узлов нечетно, то при любых α_k и β_k существует единственный (0, 2)-интерполяционный многочлен вида

$$a_0+\sum_{j=1}^{n-1}(a_j\cos jx+b_j\sin jx)+a_n\cos nx. \tag{3}$$

Фундаментальные многочлены интерполирования имеют вид

$$u_k(x)=\frac{1}{n}\left[1+\frac{2}{n}\sum_{j=1}^{n-1}\frac{(n-j)^2}{n-2j}\cos j(x-x_k)\right], \tag{4}$$

$$v_k(x)=\frac{2}{n^2}\left[\frac{1}{2n}+\sum_{j=1}^{n-1}\frac{1}{n-2j}\cos j(x-x_k)-\frac{1}{2n}\cos n(x-x_k)\right]. \tag{5}$$

Теорема 3. Если непрерывная периодическая функция $f(x)$ удовлетворяет условию

$$f(x-h)-2f(x)+f(x+h)=o(h) \tag{6}$$

и

$$\beta_{k,n}=o(n) \quad (k=0, 1, \ldots, n-1),$$

[4] J. Balázs and P. Turán, Notes on interpolation. III, *Acta Math. Acad. Sci. Hung.*, 9 (1958), стр. 195—217.

[5] G. Freud, Bemerkung über die Konvergenz eines Interpolationsverfahrens von P. Turán, *Acta Math. Acad. Sci. Hung.*, 9 (1958), стр. 337—341.

то последовательность (0, 2)-интерполяционных многочленов

$$\sum_{k=0}^{n-1}[f(x_{k,n})u_{k,n}(x)+\beta_{k,n}v_{k,n}(x)]$$

равномерно сходится к $f(x)$ на всей вещественной оси. Условие (6) не может быть заменено условием Липшица с показателем $\alpha<1$, даже если все $\beta_{k,n}$ равны нулю.

§ 2

Доказательство теоремы 1. Запишем условия

$$\left.\begin{aligned} R_n(x_k)&=\alpha_k\\ R_n''(x_k)&=\beta_k\end{aligned}\right\}\quad (k=0,1,\ldots,n-1), \tag{7}$$

которым должен удовлетворять (0, 2)-интерполяционный тригонометрический многочлен

$$R_n(x)=a_0+\sum_{j=1}^{n}(a_j\cos jx+b_j\sin jx),$$

в развернутом виде:

$$\left.\begin{aligned} a_0+\sum_{j=1}^{n}(a_j\cos jx_k+b_j\sin jx_k)&=\alpha_k\\ -\sum_{j=1}^{n}j^2(a_j\cos jx_k+b_j\sin jx_k)&=\beta_k\end{aligned}\right\}\quad (k=0,1,\ldots,n-1).$$

Если n четно, то в этих уравнениях коэффициенты при $b_{n/2}$ и b_n равны нулю, так как

$$\sin\frac{n}{2}x_k=\sin\frac{n}{2}\frac{2\pi}{n}k=\sin\pi k=0, \tag{8}$$

$$\sin nx_k=\sin 2\pi k=0.$$

Поэтому остальные $2n-1$ неизвестных a_k и b_k должны удовлетворять этой системе из $2n$ уравнений. Но это возможно лишь для исключительных значений α_k и β_k. Следовательно, условия (7), вообще говоря, не удовлетворяются, что и требовалось доказать.

§ 3

Во всем дальнейшем изложении, за исключением леммы 9, будем считать $n = 2m + 1$ нечетным натуральным числом.

Доказательству второй теоремы предпошлем четыре вспомогательных предложения. Чтобы упростить выкладки, мы сначала изучим поведение функций

$$F(x) = 1 + \frac{2}{n} \sum_{j=1}^{n-1} \frac{(n-j)^2}{n-2j} \cos jx, \tag{9}$$

$$G(x) = \frac{1}{2n} + \sum_{j=1}^{n-1} \frac{1}{n-2j} \cos jx - \frac{1}{2n} \cos nx, \tag{10}$$

с помощью которых фигурирующие в теореме 2 выражения $u_k(x)$ и $v_k(x)$, очевидно, могут быть представлены в виде

$$\left.\begin{aligned} u_k(x) &= \frac{1}{n} F(x - x_k) \qquad (11) \\ v_k(x) &= \frac{2}{n^2} G(x - x_k) \qquad (12) \end{aligned}\right\} \quad (k = 0, 1, \ldots, n-1).$$

Лемма 1. Справедливо тождество

$$F(x) = \frac{\sin n \frac{x}{2}}{\sin \frac{x}{2}} + \frac{1}{n} \sin n \frac{x}{2} \int_0^{\frac{x}{2}} \frac{n \sin t - \sin nt}{\sin^3 t} \cos t \, dt. \tag{13}$$

Доказательство. Как известно, если $n = 2m + 1$, то

$$\frac{\sin n \frac{x}{2}}{\sin \frac{x}{2}} = 1 + 2 \sum_{j=1}^{m} \cos jx. \tag{14}$$

Приступая к преобразованию второго слагаемого правой части (13), прежде всего введем обозначение

$$C(t) = 4 \sum_{j=0}^{m-1} (m-j)^2 \cos (2j+1) t \tag{15}$$

и докажем тождество

$$C(t) = \frac{n \sin t - \sin nt}{\sin^3 t} \cos t. \tag{16}$$

Так как

$$2\sin t\cos(2j+1)t=\sin(2j+2)t-\sin 2jt, \tag{17}$$

то

$$\frac{1}{2}C(t)\sin t=\sum_{j=0}^{m-1}(m-j)^2[\sin(2j+2)t-\sin 2t]=$$

$$=\sin 2mt+\sum_{j=1}^{m-1}[(m-j+1)^2-(m-j)^2]\sin 2jt=$$

$$=\sum_{j=1}^{m}[2(m-j)+1]\sin 2jt.$$

Используя равенство

$$2\sin t\sin 2jt=\cos(2j-1)t-\cos(2j+1)t,$$

получаем

$$C(t)\sin^2 t=\sum_{j=1}^{m}[2(m-j)+1][\cos(2j-1)t-\cos(2j+1)t]=$$

$$=(2m-1)\cos t-2\sum_{j=1}^{m-1}\cos(2j+1)t-\cos(2m+1)t.$$

Отсюда и из (17) следует доказываемое тождество (16):

$$2C(t)\sin^3 t=(2m-1)\sin 2t-2\sum_{j=1}^{m-1}[\sin(2j+2)t-\sin 2jt]-$$

$$-[\sin(2m+2)t-\sin 2mt]=(2m+1)\sin 2t-\sin 2mt-\sin(2m+2)t=$$

$$=2(2m+1)\sin t\cos t-2\sin(2m+1)t\cos t=2(n\sin t-\sin nt)\cos t.$$

Продолжая преобразование второго слагаемого правой части (13), получаем из (15)

$$\int_0^{\frac{x}{2}}C(t)dt=4\sum_{j=0}^{m-1}\frac{(m-j)^2}{2j+1}\sin(2j+1)\frac{x}{2}.$$

Так как

$$2\sin(2m+1)\frac{x}{2}\sin(2j+1)\frac{x}{2}=\cos(m-j)x-\cos(m+j+1)x, \tag{18}$$

то

$$\sin n\frac{x}{2}\int_0^{\frac{x}{2}}C(t)dt=2\sum_{j=0}^{m-1}\frac{(m-j)^2}{2j+1}[\cos(m-j)x-\cos(m+j+1)x]=$$

$$=2\left[\sum_{j=1}^{m}\frac{j^2}{n-2j}\cos jx+\sum_{j=m+1}^{2m}\frac{(n-j)^2}{n-2j}\cos jx\right]. \tag{19}$$

Очевидно,

$$1+\frac{1}{n}\frac{j^2}{n-2j}=\frac{n^2-2jn+j^2}{n(n-2j)}=\frac{(n-j)^2}{n(n-2j)}. \tag{20}$$

Из (14), (19) и (20) следует

$$\frac{\sin n\frac{x}{2}}{\sin\frac{x}{2}}+\frac{1}{n}\sin n\frac{x}{2}\int_0^{\frac{x}{2}}C(t)dt=1+\frac{2}{n}\sum_{j=1}^{2m}\frac{(n-j)^2}{n-2j}\cos jx.$$

Сопоставляя это равенство с (16) и (9), получаем доказываемое тождество (13).

Лемма 2. Справедливо тождество

$$G(x)=\sin n\frac{x}{2}\int_0^{\frac{x}{2}}\sin nt\operatorname{ctg}t\,dt. \tag{21}$$

Доказательство. Так как

$$\sin nt\cos t=\frac{1}{2}[\sin(n-1)t+\sin(n+1)t]$$

и для $l=2\lambda$

$$\frac{\sin lt}{\sin t}=2\sum_{j=0}^{\lambda-1}\cos(2j+1)t,$$

то при нечетном n

$$\sin nt\operatorname{ctg}t=\frac{\sin nt\cos t}{\sin t}=\frac{1}{2}\left[\frac{\sin(n-1)t}{\sin t}+\frac{\sin(n+1)t}{\sin t}\right]=$$

$$=\sum_{j=0}^{m-1}\cos(2j+1)t+\sum_{j=0}^{m+1}\cos(2j+1)t=2\sum_{j=0}^{m-1}\cos(2j+1)t+\cos nt.$$

Следовательно,

$$\int_0^{\frac{x}{2}}\sin nt\operatorname{ctg}t\,dt=2\sum_{j=0}^{m-1}\frac{1}{2j+1}\sin(2j+1)\frac{x}{2}+\frac{1}{n}\sin n\frac{x}{2}.$$

Отсюда, из (10) и (18) получаем доказываемое тождество (21):

$$\sin n\frac{x}{2}\int_0^{\frac{x}{2}}\sin nt\operatorname{ctg}t\,dt=\sum_{j=0}^{m-1}\frac{1}{2j+1}[\cos(m-j)x-\cos(m+j+1)x]+$$

$$+\frac{1}{2n}[1-\cos nx]=\frac{1}{2n}+\sum_{j=1}^{n-1}\frac{1}{n-2j}\cos jx-\frac{1}{2n}\cos nx.$$

Лемма 3. Функция $F(x)$ удовлетворяет условиям

(22) $$F(x_k)=\begin{cases} n, & \text{если} \quad k=0,\\ 0, & \text{если} \quad 0<|k|<n,\end{cases}$$

(23) $$F''(x_k)=0 \qquad (0\leqq |k|<n).$$

Доказательство. Чтобы убедиться в справедливости (22), достаточно заметить, что, как мы видели выше,

$$F(x)=\frac{\sin n\frac{x}{2}}{\sin\frac{x}{2}}+\frac{1}{n}\sin n\frac{x}{2}\int_0^{\frac{x}{2}}\frac{n\sin t-\sin nt}{\sin^3 t}\cos t\,dt,$$

(8) $$\sin\frac{n}{2}x_k=0$$

и

$$\lim_{x=0}\frac{\sin n\frac{x}{2}}{\sin\frac{x}{2}}=n.$$

Приступая к доказательству (23), для упрощения выкладок положим

(16) $$C(t)=\frac{n\sin t-\sin nt}{\sin^3 t}\cos t,$$

(24) $$H(x)=\frac{1}{\sin\frac{x}{2}}+\frac{1}{n}\int_0^{\frac{x}{2}}C(t)dt.$$

Дифференцируя тождество

$$F(x)=H(x)\sin n\frac{x}{2}$$

по правилу Лейбница, получаем $\left(\text{если } \frac{x}{2\pi} \text{ не целое число}\right)$

$$F''(x)=H''(x)\sin n\frac{x}{2}+n\,H'(x)\cos n\frac{x}{2}-\frac{n^2}{4}H(x)\sin n\frac{x}{2}.$$

Отсюда, из (8) и равенства

(25) $$\cos\frac{n}{2}x_k=\cos\pi k=(-1)^{|k|}$$

следует

$$F''(x_k)=(-1)^{|k|}\, n H'(x_k) \qquad (0<|k|<n).$$

Здесь

$$H'(x_k)=-\frac{1}{2}\frac{\cos\frac{1}{2}x_k}{\sin^2\frac{1}{2}x_k}+\frac{1}{2n}C\left(\frac{1}{2}x_k\right).$$

Но в силу (16) и (8)

$$C\left(\frac{1}{2}x_k\right)=n\frac{\cos\frac{1}{2}x_k}{\sin^2\frac{1}{2}x_k}.$$

Поэтому

$$H'(x_k)=0$$

и, следовательно,

$$F''(x_k)=0 \quad (0<|k|<n).$$

Остается доказать, что

$$F''(0)=0.$$

Дважды дифференцируя равенство

$$F(x)=1+\frac{2}{n}\sum_{j=1}^{n-1}\frac{(n-j)^2}{n-2j}\cos jx$$

и полагая $x=0$, получаем

$$F''(0)=-\frac{2}{n}\sum_{j=1}^{n-1}\frac{(n-j)^2 j^2}{n-2j}.$$

Ввиду нечетности n число слагаемых стоящей справа суммы четно. Легко видеть, что $n-j$-ое слагаемое отличается от j-ого лишь знаком. Поэтому сумма равняется нулю. Следовательно,

$$F''(0)=0,$$

что и требовалось доказать.

Лемма 4. Функция $G(x)$ удовлетворяет условиям

(26) $$G(x_k)=0 \quad (0\leqq|k|<n),$$

(27) $$G''(x_k)=\begin{cases}\frac{n^2}{2}, & \text{если} \quad k=0,\\ 0, & \text{если} \quad 0<|k|<n.\end{cases}$$

Доказательство. (26) следует из (8) и

$$G(x)=\sin n\frac{x}{2}\int\limits_{0}^{\frac{x}{2}}\sin nt\operatorname{ctg}t\,dt. \tag{21}$$

Приступая к доказательству (27), положим

$$c(t)=\sin nt\operatorname{ctg}t,$$

$$I(x)=\int\limits_{0}^{\frac{x}{2}}c(t)\,dt.$$

Дифференцируя тождество

$$G(x)=I(x)\sin n\frac{x}{2}$$

по правилу Лейбница, получаем

$$G''(x)=I''(x)\sin n\frac{x}{2}+n\,I'(x)\cos n\frac{x}{2}-\frac{n^2}{4}I(x)\sin n\frac{x}{2}.$$

Отсюда, из (8), (25) и определения $I(x)$ следует

$$G''(x_k)=(-1)^{|k|}n\,I'(x_k)=(-1)^{|k|}\frac{n}{2}c\left(\frac{1}{2}x_k\right).$$

Но

$$c\left(\frac{1}{2}x_k\right)=\sin\frac{n}{2}x_k\operatorname{ctg}x_k=0,$$

если $0<|k|<n$, и

$$c(0)=\lim_{t=0}\sin nt\operatorname{ctg}t=n.$$

Доказываемое равенство (27) является следствием трех последних соотношений.

Доказательство теоремы 2. Так как

$$\cos j(x-x_k)=\cos jx\cos jx_k+\sin jx\sin jx_k,$$

$$\cos n(x-x_k)=\cos nx,$$

то функции

$$u_k(x)=\frac{1}{n}\left[1+\frac{2}{n}\sum_{j=1}^{n-1}\frac{(n-j)^2}{n-2j}\cos j(x-x_k)\right],$$

$$v_k(x)=\frac{2}{n^2}\left[\frac{1}{2n}+\sum_{j=1}^{n-1}\frac{1}{n-2j}\cos j(x-x_k)-\frac{1}{2n}\cos n(x-x_k)\right],$$

фигурирующие в теореме 2, являются тригонометрическими многочленами вида

$$a_0+\sum_{k=1}^{n-1}(a_k\cos kx+b_k\sin kx)+a_n\cos nx. \tag{3}$$

Как мы уже отметили,

$$u_k(x)=\frac{1}{n}F(x-x_k),$$

$$v_k(x)=\frac{2}{n^2}G(x-x_k).$$

Принимая во внимание (22), (23), (26) и (27), получаем отсюда

$$u_k(x_j)=\frac{1}{n}F(x_{j-k})=\begin{cases}1, & \text{если} \quad j=k,\\ 0, & \text{если} \quad j\neq k,\end{cases}$$

$$\left.\begin{aligned} u_k''(x_j)&=\frac{1}{n}F''(x_{j-k})=0,\\ v_k(x_j)&=\frac{2}{n^2}G(x_{j-k})=0,\end{aligned}\right. \qquad (j,k=0,1,\ldots,n-1),$$

$$v_k''(x_j)=\frac{2}{n^2}G''(x_{j-k})=\begin{cases}1, & \text{если} \quad j=k,\\ 0, & \text{если} \quad j\neq k.\end{cases}$$

Из сказанного следует, что функции $u_k(x)$ и $v_k(x)$ действительно являются фундаментальными многочленами (0, 2)-интерполирования вида (3).

Тригонометрический многочлен вида (3), удовлетворяющий условиям

$$\left.\begin{aligned} R_n(x_k)&=\alpha_k\\ R_n''(x_k)&=\beta_k\end{aligned}\right\} \qquad (k=0,1,\ldots,n-1), \tag{7}$$

очевидно, может быть записан в форме

$$R_n(x)=\sum_{k=0}^{n-1}[\alpha_k u_k(x)+\beta_k v_k(x)].$$

Следовательно, (0, 2)-интерполяционный многочлен вида (3) существует при любом выборе чисел α_k и β_k.

Из существования этих многочленов следует и их единственность. Чтобы доказать этот факт, перепишем уравнения (7) в виде

$$\left.\begin{aligned} a_0+\sum_{j=1}^{n-1}(a_j\cos jx_k+b_j\sin jx_k)+a_n\cos nx_k&=\alpha_k\\ -\sum_{j=1}^{n-1}j^2(a_j\cos jx_k+b_j\sin jx_k)-n^2a_n\cos nx_k&=\beta_k\end{aligned}\right\} \quad (k=0,1,\ldots,n-1).$$

Эта система уравнений разрешима при любых α_k и β_k в том и только в том случае, если ее решение единственно, что мы и хотели доказать.

Доказательство теоремы 2 завершено.

§ 4

Доказательству третьей теоремы мы также предпошлем несколько лемм.

Лемма 5. Имеет место неравенство

$$|F(x)|<5n. \tag{28}$$

Доказательство. Из доказанного выше тождества

$$F(x)=\frac{\sin n\frac{x}{2}}{\sin\frac{x}{2}}+\frac{1}{n}\sin n\frac{x}{2}\int_0^{\frac{x}{2}}\frac{n\sin t-\sin nt}{\sin^3 t}\cos t\,dt \tag{13}$$

и неравенства

$$|\sin nt|\leqq n|\sin t|$$

получаем

$$|F(x)|<n+\frac{1}{n}\int_0^{\frac{\pi}{2}}\frac{n\sin t-\sin nt}{\sin^3 t}\,dt\qquad(0\leqq x\leqq\pi). \tag{29}$$

Чтобы оценить стоящий здесь интеграл, разобьем его на два, распространенные на отрезки $\left(0,\frac{\pi}{2n}\right)$ и $\left(\frac{\pi}{2n},\frac{\pi}{2}\right)$, соответственно. Так как при $0<x<\frac{\pi}{2}$

$$x>\sin x>\begin{cases}\frac{2}{\pi}x\\ x-\frac{x^3}{6}\end{cases},$$

то

$$\int_0^{\frac{\pi}{2n}}\frac{n\sin t-\sin nt}{\sin^3 t}\,dt<\int_0^{\frac{\pi}{2n}}\frac{nt-\left(nt-\frac{n^3t^3}{6}\right)}{\left(\frac{2}{\pi}t\right)^3}\,dt=\frac{\pi^4}{96}n^2.$$

С другой стороны,

$$\int\limits_{\frac{\pi}{2n}}^{\frac{\pi}{2}} \frac{n\sin t - \sin nt}{\sin^3 t}\,dt < \int\limits_{\frac{\pi}{2n}}^{\frac{\pi}{2}} \left(\frac{n}{\sin^2 t} + \frac{1}{\sin^3 t}\right) dt <$$

$$< \int\limits_{\frac{\pi}{2n}}^{+\infty} \left(\frac{\pi^2}{4} n \frac{1}{t^2} + \frac{\pi^3}{8}\frac{1}{t^3}\right) dt = \frac{\pi}{2} n^2 + \frac{\pi}{4} n^2 = \frac{3\pi}{4} n^2.$$

Поэтому

$$\int\limits_{0}^{\frac{\pi}{2}} \frac{n\sin t - \sin nt}{\sin^3 t}\,dt < \left(\frac{\pi^4}{96} + \frac{3\pi}{4}\right) n^2 < 4n^2.$$

Отсюда и из (29) получаем (28) для $0 \leqq x \leqq \pi$. Но $F(x)$ четный тригонометрический многочлен. Поэтому (28) выполняется для всех x, что и требовалось доказать.

Лемма 6. Справедливо неравенство

$$\left| F\left[(2j+1)\frac{\pi}{n}\right] \right| > \frac{n}{5} \qquad (j = 0, \pm 1, \pm 2, \ldots). \tag{30}$$

Доказательство. Из (13) следует

$$F\left[(2j+1)\frac{\pi}{n}\right] = (-1)^{|j|}\left[\frac{1}{\sin\left(j+\frac{1}{2}\right)\frac{\pi}{n}} + \frac{1}{n}\int\limits_{0}^{\left(j+\frac{1}{2}\right)\frac{\pi}{k}} \frac{n\sin t - \sin nt}{\sin^3 t}\cos t\,dt\right].$$

Поэтому

$$\left| F\left[(2j+1)\frac{\pi}{n}\right] \right| \geqq 1 + \frac{1}{n}\int\limits_{0}^{\frac{\pi}{2n}} \frac{n\sin t - \sin nt}{\sin^3 t}\cos t\,dt \quad (j = 0, 1, \ldots, m).$$

Полагая $n \geqq 3$ (при $n = 1$ (30) очевидно) и принимая во внимание неравенства

$$x > x - \frac{x^3}{6} + \frac{x^5}{120} > \sin x > x - \frac{x^3}{6} \qquad \left(0 < x < \frac{\pi}{2}\right),$$

$$\cos x > \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2} \qquad \left(0 < x < \frac{\pi}{6}\right),$$

получаем отсюда

$$\left|F\left[(2j+1)\frac{\pi}{n}\right]\right|>1+\frac{\sqrt{3}}{2}\frac{1}{n}\int_0^{\frac{\pi}{2n}}\frac{n\left(t-\frac{t^3}{6}\right)-\left(nt-\frac{n^3t^3}{6}+\frac{n^5t^5}{120}\right)}{t^3}dt=$$

$$=1+\frac{\sqrt{3}}{2}\int_0^{\frac{\pi}{2n}}\left(\frac{n^2-1}{6}-\frac{n^4}{120}t^2\right)dt=1+\frac{\sqrt{3}}{2}\left[\frac{\pi}{12}n-\frac{\pi}{12}\frac{1}{n}-\frac{\pi^3}{2880}n\right]>$$

$$>\frac{\sqrt{3}}{2}\left[\frac{\pi}{12}-\frac{\pi^3}{2880}\right]n>\frac{n}{5}\quad (j=0, 1, \ldots, m).$$

Ввиду того, что $F(x)$ есть четный тригонометрический многочлен, это неравенство верно при любых j, что и требовалось доказать.

Лемма 7. Имеет место неравенство

$$|G(x)|<\pi. \tag{31}$$

Доказательство. Пусть

$$0<x\leqq\pi,\quad \frac{2\pi}{n}k<x\leqq\frac{2\pi}{n}(k+1).$$

Очевидно

$$\int_0^{\frac{x}{2}}\sin nt\operatorname{ctg}t\,dt=\sum_{l=0}^{k-1}\int_{l\frac{\pi}{n}}^{(l+1)\frac{\pi}{n}}\sin nt\operatorname{ctg}t\,dt+\int_{k\frac{\pi}{n}}^{\frac{x}{2}}\sin nt\operatorname{ctg}t\,dt.$$

Так как в стоящем справа $l+1$-ом интеграле I_l

$$\operatorname{ctg}t>0,\quad \operatorname{sign}\sin nt=(-1)^l,$$

то

$$\operatorname{sign}I_l=(-1)^l\qquad (l=0, 1, \ldots, k). \tag{32}$$

Ввиду того, что функция $\operatorname{ctg}t$ строго убывающая, а $|\sin nt|$ $\frac{\pi}{n}$-периодична, справедливы неравенства

$$|I_0|>|I_1|>\cdots>|I_k|. \tag{33}$$

Поэтому

$$0<\int_0^{\frac{x}{2}}\sin nt\operatorname{ctg}t\,dt<I_0.$$

Но в I_0

$$0<\sin nt<n\sin t,\quad 0<\cos t<1.$$

(Можно считать $n \geqq 3$, так как при $n=1$ (31) очевидно.) Следовательно,

$$I_0 = \int_0^{\frac{\pi}{n}} \sin nt \operatorname{ctg} t\, dt = \int_0^{\frac{\pi}{n}} \frac{\sin nt}{\sin t} \cos t\, dt < \int_0^{\frac{\pi}{n}} n\, dt = \pi.$$

Поэтому

$$0 < \int_0^{\frac{x}{2}} \sin nt \operatorname{ctg} t\, dt < \pi.$$

Так как

$$G(x) = \sin n \frac{x}{2} \int_0^{\frac{x}{2}} \sin nt \operatorname{ctg} t\, dt, \tag{21}$$

то (31) имеет место для $0 < x \leqq \pi$. Но $G(x)$ есть четный тригонометрический многочлен. Поэтому (31) должно выполняться для всех x, что и требовалось доказать.

Лемма 8. Справедливо неравенство

$$\left| G\left[(2j+1)\frac{\pi}{n}\right] \right| > \frac{1}{2} \qquad (j = 0, \pm 1, \pm 2, \ldots). \tag{34}$$

Доказательство. Из (21) получаем

$$G\left[(2j+1)\frac{\pi}{n}\right] = (-1)^{|j|} \int_0^{\left(j+\frac{1}{2}\right)\frac{\pi}{n}} \sin nt \operatorname{ctg} t\, dt. \tag{35}$$

Так как

$$\int_0^{\left(j+\frac{1}{2}\right)\frac{\pi}{n}} \sin nt \operatorname{ctg} t\, dt = \sum_{l=0}^{j-1} \int_{l\frac{\pi}{n}}^{(l+1)\frac{\pi}{n}} \sin nt \operatorname{ctg} t\, dt + \int_{j\frac{\pi}{n}}^{\left(j+\frac{1}{2}\right)\frac{\pi}{n}} \sin nt \operatorname{ctg} t\, dt$$

и для стоящих справа интегралов I_l имеет место (32) и (33), то

$$\int_0^{\left(j+\frac{1}{2}\right)\frac{\pi}{n}} \sin nt \operatorname{ctg} t\, dt > I_0 + I_1 = \int_0^{\frac{2\pi}{n}} \sin nt \operatorname{ctg} t\, dt \qquad (0 < j \leqq m).$$

Оценим стоящий справа интеграл.

$$\int_0^{\frac{2\pi}{n}} \sin nt \operatorname{ctg} t\, dt = \int_0^{\frac{\pi}{n}} \sin nt \left[\operatorname{ctg} t - \operatorname{ctg}\left(t + \frac{\pi}{n}\right)\right] dt >$$

$$> \int_0^{\frac{\pi}{2n}} \sin nt \left[\operatorname{ctg} t - \operatorname{ctg}\left(t + \frac{\pi}{n}\right)\right] dt.$$

(Мы считаем, что $n \geqq 3$. При $n=1$ (34) очевидно.) Но

$$\operatorname{ctg} t-\operatorname{ctg}\left(t+\frac{\pi}{n}\right)=\frac{\cos t}{\sin t}-\frac{\cos\left(t+\frac{\pi}{n}\right)}{\sin\left(t+\frac{\pi}{n}\right)}=$$

$$=\frac{\cos t\sin\left(t+\frac{\pi}{n}\right)-\sin t\cos\left(t+\frac{\pi}{n}\right)}{\sin t\sin\left(t+\frac{\pi}{n}\right)}=\frac{\sin\frac{\pi}{n}}{\sin t\sin\left(t+\frac{\pi}{n}\right)}>\frac{\sin\frac{\pi}{n}}{\sin t\sin 2\frac{\pi}{n}}=$$

$$=\frac{1}{2\sin t\cos\frac{\pi}{n}}>\frac{1}{2\sin t}.$$

Поэтому

$$\int_0^{\frac{2\pi}{n}}\sin nt\operatorname{ctg} t\,dt>\frac{1}{2}\int_0^{\frac{\pi}{2n}}\frac{\sin nt}{\sin t}dt>\frac{1}{2}\int_0^{\frac{\pi}{2n}}\frac{\frac{2}{\pi}nt}{t}dt=\frac{1}{2}.$$

Следовательно, при $0<j\leqq m$

$$\int_0^{\left(j+\frac{1}{2}\right)\frac{\pi}{n}}\sin nt\operatorname{ctg} t\,dt>\frac{1}{2}. \tag{36}$$

Это неравенство верно и для $j=0$, так как

$$\int_0^{\frac{\pi}{2n}}\sin nt\operatorname{ctg} t\,dt=\int_0^{\frac{\pi}{2n}}\frac{\sin nt}{\sin t}\cos t\,dt>$$

$$>\int_0^{\frac{\pi}{2n}}\frac{2}{\pi}n\cos\frac{2\pi}{n}dt>\frac{\sqrt{3}}{2}>\frac{1}{2}.$$

Из (35) и (36) (34) следует при $j=0, 1, \ldots, m$. Ввиду четности и 2π-периодичности $G(x)$ это неравенство должно выполняться для любых j, что и требовалось доказать.

Лемма 9. Пусть непрерывная периодическая функция $f(x)$ удовлетворяет условию

$$f(x-h)-2f(x)+f(x+h)=o(h). \tag{6}$$

Существует последовательность тригонометрических многочленов $T_n(x)$ не выше $n-1$-ого порядка, для которых

$$f(x)-T_n(x)=o\left(\frac{1}{n}\right), \tag{37}$$

$$T_n''(x)=o(n). \tag{38}$$

Доказательство. А. Зигмунд доказал,[6] что, если функция $f(x)$ удовлетворяет условию леммы, то существуют тригонометрические многочлены $U_n(x)$ не выше $n-1$-ого порядка, для которых

$$f(x)-U_n(x)=o\left(\frac{1}{n}\right).$$

Лемма 9 является следствием этой теоремы. Нижеследующее доказательство этого факта аналогично соответствующему доказательству статьи[5]. Его идея принадлежит П. Турану.

Пусть

$$2^k \leqq n < 2^{k+1}.$$

Положим

$$T_n(x)=U_{2^k}(x).$$

Выполнение условия (37) очевидно.

Чтобы доказать (38), заметим, что

$$T_n(x)=\sum_{j=1}^{k}[U_{2^j}(x)-U_{2^{j-1}}(x)]+U_1(x).$$

Из теоремы А. Зигмунда следует соотношение

$$U_{2^j}(x)-U_{2^{j-1}}(x)=o\left(\frac{1}{2^j}\right).$$

Но тогда в силу неравенства С. Н. Бернштейна

$$U_{2^j}''(x)-U_{2^{j-1}}''(x)=o(2^j),$$

и поэтому

$$T_n''(x)=\sum_{j=1}^{k}[U_{2^j}''(x)-U_{2^{j-1}}''(x)]=o(2^{k+1})=o(n),$$

что и требовалось доказать.

Доказательство теоремы 3. Пусть непрерывная периодическая функция $f(x)$ удовлетворяет условию (6). Введем обозначение

$$R_n(x)=\sum_{k=0}^{n-1}[f(x_k)u_k(x)+\beta_k v_k(x)].$$

[6] A. Zygmund, Smooth functions, *Duke Math. Journal*, 12 (1945), стр. 47—76.

(Чтобы упростить запись, мы опускаем индекс n в величинах $x_{k,n}$ $\beta_{k,n}$, $u_{k,n}(x)$, $v_{k,n}(x)$.) Для фигурирующего в предыдущей лемме тригонометрического многочлена $T_n(x)$ (как и для всякого многочлена вида (3)), очевидно, выполняется тождество

$$T_n(x)=\sum_{k=0}^{n-1}[T_n(x_k)u_k(x)+T_n''(x_k)v_k(x)].$$

Поэтому

$$f(x)-R_n(x)=f(x)-T_n(x)+T_n(x)-R_n(x)=f(x)-T_n(x)+$$
$$+\sum_{k=0}^{n-1}[T_n(x_k)u_k(x)+T_n''(x_k)v_k(x)]-\sum_{k=0}^{n-1}[f(x_k)u_k(x)+\beta_k v_k(x)]=$$
$$=f(x)-T_n(x)+\sum_{k=0}^{n-1}[T_n(x_k)-f(x_k)]u_k(x)+\sum_{k=0}^{n-1}[T_n''(x_k)-\beta_k]v_k(x).$$

Выше мы видели, что

$$f(x)-T_n(x)=o\left(\frac{1}{n}\right),$$
$$T_n''(x)=o(n).$$

Будем считать, что

$$\beta_k=o(n).$$

Тогда

$$f(x)-R_n(x)=o\left(\frac{1}{n}\right)+o\left(\frac{1}{n}\right)\sum_{k=0}^{n-1}|U_k(x)|+o(n)\sum_{k=0}^{n-1}|v_k(x)|. \tag{39}$$

Из ранее доказанных равенств

$$u_k(x)=\frac{1}{n}F(x-x_k), \tag{11}$$

$$v_k(x)=\frac{2}{n^2}G(x-x_k) \tag{12}$$

и неравенств

$$|F(x)|<5n,\quad |G(x)|<\pi$$

получаем

$$|u_k(x)|<5,\quad |v_k(x)|<\frac{2\pi}{n^2}.$$

Поэтому

$$\sum_{k=0}^{n-1}|u_k(x)|<5n,\quad \sum_{k=0}^{n-1}|v_k(x)|<\frac{2\pi}{n}.$$

Отсюда и из (39) следует равномерная сходимость последовательности $R_n(x)$ к $f(x)$. Таким образом, первая часть теоремы 3 доказана.

В леммах 6 и 8 было доказано, что

$$\left|F\left[(2j+1)\frac{\pi}{n}\right]\right| > \frac{n}{5}, \quad \left|G\left[(2j+1)\frac{\pi}{n}\right]\right| > \frac{1}{2}.$$

Из этих неравенств и равенств (11) и (12) получаем

$$\left|u_k\left[(2j+1)\frac{\pi}{n}\right]\right| > \frac{1}{5}, \quad \left|v_k\left[(2j+1)\frac{\pi}{n}\right]\right| > \frac{1}{n^2}.$$

Полагая $j=m$ и суммируя, приходим к оценкам

$$\sum_{k=0}^{n-1} |u_k(\pi)| > \frac{n}{5}, \quad \sum_{k=0}^{n-1} |v_k(\pi)| > \frac{1}{n}. \tag{40}$$

Первая из них понадобится нам в дальнейшем. Вторая показывает, что условие

$$\beta_k = o(n)$$

необходимо.

Нам остается доказать, что условие А. Зигмунда

$$f(x-h) - 2f(x) + f(x+h) = o(h)$$

не может быть заменено условием Липшица с показателем $1-\varepsilon$ ни при каком ε $(0<\varepsilon<1)$.

Это можно сделать также, как доказывается аналогичная теорема в работе.[7]

Пусть непрерывная функция $\sum\limits_{k=0}^{n-1} |u_k(x)|$ принимает свое наибольшее значение в точке ξ_n, т. е.

$$\lambda_n = \sum_{k=0}^{n-1} |u_k(\xi_k)|, \tag{41}$$

$$\sum_{k=0}^{n-1} |u_k(x)| \leqq \lambda_n. \tag{42}$$

Неравенство (40) может быть переписано в виде

$$\lambda_n > \frac{n}{5}. \tag{43}$$

Обозначим через $\varphi_n(x)$ функцию, линейную между каждой соседней парой узлов x_k и принимающую в этих точках значения

$$\varphi_n(x_k) = \operatorname{sign} u_k(\xi_n) \qquad (k=0, 1, \ldots, n). \tag{44}$$

[7] P. Erdős and P. Turán, On the role of the Lebesgue functions in the theory o Lagrange interpolation, *Acta Math. Acad. Sci. Hung.*, 6 (1955), стр. 47—66.

Легко видеть, что

$$(45)\qquad |\varphi_n(x)| \leqq 1, \quad |\varphi_n(x+h)-\varphi_n(x)| \leqq \begin{cases} \dfrac{n}{\pi}h \\ 2 \end{cases}.$$

Поэтому эта функция удовлетворяет условию Липшица с показателем 1 и для любого ε $(0<\varepsilon<1)$

$$(46)\qquad |\varphi_n(x+h)-\varphi_n(x)| \leqq \left(\frac{n}{\pi}h\right)^{1-\varepsilon} 2^{\varepsilon} < 2n^{1-\varepsilon}h^{1-\varepsilon}.$$

Введем еще обозначение

$$R_n[f, x] = \sum_{k=0}^{n-1} f(x_k)u_k(x).$$

Принимая во внимание (44) и (41), получаем отсюда

$$(47)\qquad R_n[\varphi_n, \xi_n] = \sum_{k=0}^{n-1} \varphi_n(x_k)u_k(\xi_n) = \sum_{k=0}^{n-1} |u_k(\xi_n)| = \lambda_n.$$

С другой стороны, при любых n и p

$$(48)\qquad |R_n[\varphi_p, x]| = \left|\sum_{k=0}^{n-1} \varphi_p(x_k)u_k(x)\right| \leqq \sum_{k=0}^{n-1} |u_k(x)| \leqq \lambda_n.$$

Здесь было использовано (45) и (42).

Так как при любом p $\varphi_p(x)$ удовлетворяет условию Липшица с показателем 1, то можно считать, что последовательность $R_n[\varphi_p, x]$ равномерно сходится к $\varphi_p(x)$. (В противном случае теорема доказана.) Отсюда и из неравенства

$$|\varphi_p(x)| \leqq 1$$

следует существование таких N_p, что

$$(49)\qquad |R_n[\varphi_p, x]| < 2, \quad \text{если} \quad n > N_p.$$

Пусть последовательность нечетных чисел n_i удовлетворяет условиям

$$(50)\qquad n_1 > 6, \quad n_{i+1} > 6n_i,$$

$$(51)\qquad n_{i+1} > N_{n_i} \qquad (i = 1, 2, 3, \ldots).$$

Положим

$$0 < \varepsilon < 1$$

и выберем q так, чтобы выполнялось неравенство

$$(52)\qquad 1 < q < 2^{\varepsilon} < 2.$$

В силу (50) и (52)

(53) $$\frac{q^i}{n_i} < \frac{2^i}{6^i} = \frac{1}{3^i}.$$

Отсюда и из (45) следует, что бесконечный ряд

$$\sum_{i=1}^{\infty} \frac{q^i}{n_i} \varphi_{n_i}(x)$$

равномерно сходится. Обозначим его сумму через $f(x)$ и докажем, что эта функция удовлетворяет условию Липшица с показателем $1-\varepsilon$.

Принимая во внимание (46), (50) и (52), видим, что

$$|f(x+h)-f(x)| = \left|\sum_{i=1}^{\infty} \frac{q^i}{n^i} [\varphi_{n_i}(x+h)-\varphi_{n_i}(x)]\right| \leqq$$

$$\leqq \sum_{i=1}^{\infty} \frac{q^i}{n^i} |\varphi_{n_i}(x+h)-\varphi_{n_i}(x)| < 2h^{1-\varepsilon} \sum_{i=1}^{\infty} \frac{q^i}{n_i^{\varepsilon}} <$$

$$< 2h^{1-\varepsilon} \sum_{i=1}^{\infty} \frac{(2^{\varepsilon})^i}{(6^i)^{\varepsilon}} = \frac{2}{3^{\varepsilon}-1} h^{1-\varepsilon},$$

что и требовалось доказать.

Докажем, наконец, что

(54) $$\lim_{j\to\infty} R_{n_j}[f, \xi_{n_j}] = \infty,$$

и поэтому последовательность $R_n[f, x]$ не сходится равномерно к $f(x)$.

Очевидно,

(55) $$R_{n_j}[f, x] = \sum_{i=1}^{\infty} \frac{q^i}{n^i} R_{n_j}[\varphi_{n_i}, x] =$$
$$= \sum_{i=1}^{j-1} \frac{q^i}{n^i} R_{n_j}[\varphi_{n_i}, x] + \frac{q^j}{n_j} R_{n_j}[\varphi_{n_j}, x] + \sum_{i=j+1}^{\infty} \frac{q^i}{n_i} R_{n_j}[\varphi_{n_i}, x].$$

В силу (53), (51) и (49)

(56) $$\left|\sum_{i=1}^{j-1} \frac{q^i}{n^i} R_{n_j}[\varphi_{n_i}, x]\right| < 2\sum_{i=1}^{j-1} \frac{1}{3^i} < 1.$$

С другой стороны, в силу (48), (50) и (52),

(57) $$\left|\sum_{i=j+1}^{\infty} \frac{q^i}{n^i} R_{n_j}[\varphi_{n_i}, x]\right| \leqq \lambda_{n_j} \sum_{i=j+1}^{\infty} \frac{q^i}{n^i} < \lambda_{n_j} \frac{q^j}{n_j} \sum_{i=1}^{\infty} \frac{1}{3^i} = \frac{1}{2} \lambda_{n_j} \frac{q^j}{n_j}.$$

Резюмируя (55), (56), (47) и (57), получаем

$$R_{n_j}[f, \xi_{n_j}] > \frac{q^j}{n^j} \lambda_{n_j} - \frac{1}{2} \lambda_{n_j} \frac{q^j}{n_j} - 1 = \frac{1}{2} \lambda_{n_j} \frac{q^j}{n_j} - 1.$$

Отсюда и из (43) следует

$$R_{n_j}[f, \xi_{n_j}] > \frac{1}{10} q^j - 1.$$

Так как $q > 1$, то (54) действительно имеет место.

(Поступило 24. XI. 1959.)

SOME RESULTS AND PROBLEMS ON SET THEORY

By
A. HAJNAL (Budapest)
(*Presented by* P. ERDŐS)

1. Introduction. (A short summary of the results and problems investigated in this paper.)

a) Let S be a set, and let $f(x)$ be a function which makes correspond to every $x \in S$ a subset $f(x)$ of S so that $x \notin f(x)$. We shall call such a function $f(x)$ a set-mapping defined on S. A subset $S' \subseteq S$ is called free (or independent) with respect to the set-mapping $f(x)$ if for every $x \in S'$ and $y \in S'$, $x \notin f(y)$ and $y \notin f(x)$.

In their paper [2] ERDŐS and FODOR raised the following unsolved problem: Suppose $\bar{S} = m$, $\overline{f(x)} < m$ for every $x \in S$, let $n < m$ be a cardinal number such that $\overline{f(x) \cap f(y)} < n$ for every $x, y \in S$ $(x \neq y)$. Does there then always exist a free subset of power m?[1]

In Section 2, using the generalized continuum hypothesis, we are going to prove that the answer to this question is negative, provided $m = \aleph_{\alpha+1}$, $n = \aleph_\alpha$ where $\aleph_\alpha$ is regular (see Theorem 1), and comparing our result with the positive results of ERDŐS and FODOR we are going to formulate the simplest unsolved problems. The theorems in the proof of which the generalized continuum hypothesis is used will be denoted by (*).

b) In Section 3 we are going to prove some special results concerning the partition symbol for types introduced by ERDŐS and RADO.[2]

Especially, we are interested in the question when the symbols $\lambda \to (\alpha, \beta)^2$, $\omega_1 \to (\alpha, \beta)^2$ hold where λ is the type of the set of the real numbers and α, β are ordinal numbers.

Our main results are the following:

$\omega_1 \not\to (\omega + 2, \omega_1)^2$.

$\lambda \to (\omega \cdot n, \alpha)^2$ for every finite n and for every $\alpha < \omega_1$.

$\lambda \to (\eta, \alpha \vee \alpha^*)^2$ for every $\alpha < \omega_1$ where η denotes the type of the set of the rational numbers.

$\omega_1 \to (\omega \cdot 2, \omega \cdot n)^2$ for every finite n.

(See Theorems 5, 6, 7, 8, respectively.)

[1] See [2], p. 251.
[2] See e. g. [1], p. 430, (2).

c) Let S be a set, $\overline{\overline{S}}=m$. Let p, q, r be cardinal numbers such that $m \geqq p \geqq q \geqq r$.

We raise the following general problem: Under what conditions for m, p, q, r does there exist a system $\mathfrak{S}$ of subsets of S which satisfies the following conditions:

(i) $\overline{\overline{X}}=q$ for every $X \in \mathfrak{S}$;

(ii) $\overline{\overline{X \cap Y}}<r$ for every $X \in \mathfrak{S}, Y \in \mathfrak{S}$ $(X \neq Y)$ (i. e. $\mathfrak{S}$ consists of almost disjoint sets);

(iii) for an arbitrary $Z \subseteq S$ $(\overline{\overline{Z}}=p)$ there exists an $X \in \mathfrak{S}$ such that $X \subseteq Z$.

In Section 4, using the generalized continuum hypothesis, we are going to prove that for $m=\aleph_{\alpha+1}$, $p=\aleph_{\alpha+1}$, $q=\aleph_\alpha$, $r=\aleph_\alpha$ the answer is affirmative, provided that $\aleph_\alpha$ is regular (see Theorem 9).

This is the solution of the problem in the simplest non-trivial case.

It is quite easy to see that for the case $q>r$ or $p=q$ the answer is generally negative.

Since we can not give a complete discussion, we restrict ourselves to the proof of the above-mentioned result.

2. Set-mappings.

(*) THEOREM 1. *Let S be a set, $\overline{\overline{S}}=\aleph_{\alpha+1}$. Suppose $\aleph_\alpha$ is regular.*

One can define a set-mapping $f(x)$ on S which satisfies the following conditions:

(o) $\overline{\overline{f(x)}}=\aleph_\alpha$ *for every* $x \in S$;

(oo) $\overline{\overline{f(x) \cap f(y)}}<\aleph_\alpha$ *for every* $x, y \in S$ $(x \neq y)$;

(ooo) *every subset S' of S which is free with respect to $f(x)$ is of power less than $\aleph_{\alpha+1}$.*

Instead of Theorem 1 we are going to prove the following more general

(*) THEOREM 2. *Let S be a set, $\overline{\overline{S}}=\aleph_{\alpha+1}$. Suppose $\aleph_\alpha$ is regular. Then there exists a system $\mathfrak{S}$ of subsets of S which satisfies the following conditions:*

(+) $\overline{\overline{\mathfrak{S}}}=\aleph_{\alpha+1}$;

(++) $\overline{\overline{X}}=\aleph_\alpha$ *for every* $X \in \mathfrak{S}$;

(+++) $\overline{\overline{X \cap Y}}<\aleph_\alpha$ *for every* $X, Y \in \mathfrak{S}$ $(X \neq Y)$, *i. e. the elements of $\mathfrak{S}$ are almost disjoint;*

(++++) *for an arbitrary* $\mathfrak{S}' \subseteq \mathfrak{S}$, *if* $\overline{\overline{\mathfrak{S}'}}=\aleph_{\alpha+1}$, *then*

$$\overline{\overline{S-\bigcup_{X \in \mathfrak{S}'} X}}<\aleph_{\alpha+1}.$$

PROOF. The set $[S]^{\aleph_\alpha}$ is of power $\aleph_{\alpha+1}$ by the hypothesis $(*)$.[3] Let $[S]^{\aleph_\alpha}=\{A_\nu\}_{\nu<\omega_{\alpha+1}}$ be a well-ordering of type $\omega_{\alpha+1}$ of the set $[S]^{\aleph_\alpha}$. Let $\mathfrak{A}_\nu$ denote the set $\{A_\mu\}_{\mu<\nu}$. We are going to define a sequence $\{B_\nu\}_{\nu<\omega_{\alpha+1}}$ of type $\omega_{\alpha+1}$ of the subsets of S by transfinite induction on ν as follows:

(1) Let B_0 be an arbitrary subset of power $\aleph_\alpha$ of S. Suppose that the sets B_μ are already defined for every $\mu<\nu$ where $\nu<\omega_{\alpha+1}$.

Put $\mathfrak{B}_\nu=\{B_\mu\}_{\mu<\nu}$.

We say that an element A_μ of $\mathfrak{A}_\nu$ is good at the ν^{th} step if it is not contained in the sum of less than $\aleph_\alpha$ B_μ already defined, i. e:

(2) A_μ is good at the ν^{th} step if the set $A_\mu-\bigcup\limits_{B_{\mu'}\in\mathfrak{B}} B_{\mu'}$ is non-empty for an arbitrary $\mathfrak{B}\subseteq\mathfrak{B}_\nu$, provided $\overline{\overline{\mathfrak{B}}}<\aleph_\alpha$.

(3) Let $\mathfrak{A}^*_\nu$ be the set of all A_μ's which are good at the ν^{th} step.

It is obvious that

(4) $\overline{\overline{\mathfrak{A}^*_\nu}}\leqq\aleph_\alpha$ and $\overline{\overline{\mathfrak{B}_\nu}}\leqq\aleph_\alpha$, therefore there exist well-orderings

(5) $\mathfrak{A}^*_\nu=\{A^\nu_\varrho\}_{\varrho<\varphi}$ of type φ and $\mathfrak{B}_\nu=\{B^\nu_\varrho\}_{\varrho<\varphi'}$ of type φ' of the sets $\mathfrak{A}^*_\nu$ and $\mathfrak{B}_\nu$, respectively, where $\varphi,\varphi'\leqq\omega_\alpha$.

Here we may suppose $\varphi=\varphi'=\omega_\alpha$, since if $\varphi<\omega_\alpha$ or $\varphi'<\omega_\alpha$, the definition of a B_ν satisfying the essential requirements of (6), (7), (8), (9) is trivial. Let now ϱ be an arbitrary ordinal number less than ω_α.

(6) Let x^ν_ϱ be an element of the set $A^\nu_\varrho-\bigcup\limits_{\varrho'<\varrho} B^\nu_{\varrho'}$. Such an element x^ν_ϱ exists by (2), (3) (4) and (5).

Put

(7)
$$B_\nu=\{x^\nu_\varrho\}_{\varrho<\omega_\alpha}.$$

So the B_ν's are defined for every $\nu<\omega_{\alpha+1}$. We are going to prove that $\mathfrak{S}=\{B_\nu\}_{\nu<\omega_{\alpha+1}}$ satisfies our requirements.

First of all it is obvious from (1) and (7) that

(8) $\overline{\overline{B}}_\nu\leqq\aleph_\alpha$ for every $\nu<\omega_{\alpha+1}$, i. e. $(++)$ holds.[4]

[3] $[X]^a$ denotes, as usual, the set of all subsets of X of power a.

[4] We have to write here $\leqq$, since in (7) the x^ν_ϱ's are not necessarily different. But if we can prove that $\mathfrak{S}$ satisfes the requirements of Theorem 2 with $\leqq$ instead of $=$ in $(++)$, then we may suppose that $(++)$ holds with $=$, since such a system $\mathfrak{S}$ contains only less than $\aleph_\alpha$ sets of power less than $\aleph_\alpha$ as it is shown by the lemma proved in [3], p. 55, and we can omit these sets.

The B_ν's are almost disjoint, i. e.

(9) $$\overline{\overline{B_\mu \cap B_\nu}} < \aleph_\alpha \quad \text{if} \quad \mu \neq \nu.$$

We may suppose $\mu < \nu$. Then by (5) there is a $\varrho < \omega_\alpha$ such that $B_\mu = B_\varrho^\nu$, and therefore $x_{\varrho'}^\nu \notin B_\mu$ for $\varrho' > \varrho$ by (6), hence (9) follows from (7).

It follows that the B_ν's are different, and so we have

(10) $$\overline{\overline{\mathfrak{S}}} = \aleph_{\alpha+1}.$$

Let now $\mathfrak{A}^*$ be the set of those A_μ's which are good at the ν^{th} step for every ν, i. e.

(11) $\mathfrak{A}^*$ is the set of all A_μ's for which $A_\mu - \bigcup_{B_{\mu'} \in \mathfrak{B}} B_{\mu'}$ is non-empty, provided $\mathfrak{B} \subseteq \mathfrak{S}$, $\overline{\overline{\mathfrak{B}}} < \aleph_\alpha$.

Now we are going to prove:

(12) Every subset $S' \subseteq S$ $(\overline{\overline{S'}} = \aleph_{\alpha+1})$ contains an $A_{\mu_0} \in \mathfrak{A}^*$.

Let $S' \subseteq S$, $\overline{\overline{S'}} = \aleph_{\alpha+1}$. Let $\mathfrak{B}$ be the system of those B_μ's for which $\overline{\overline{B_\mu \cap S'}} = \aleph_\alpha$.

We distinguish two cases:

(a) $\overline{\overline{\mathfrak{B}}} < \aleph_\alpha$, (b) $\overline{\overline{\mathfrak{B}}} \geqq \aleph_\alpha$.

(a) The set $S' - \bigcup_{B_\mu \in \mathfrak{B}} B_\mu$ is of power $\aleph_{\alpha+1}$, and every subset A_{μ_0} of it belongs to $\mathfrak{A}^*$, since then $\overline{\overline{A_{\mu_0} \cap B_{\mu'}}} < \aleph_\alpha$ for every $\mu' < \omega_{\alpha+1}$, and so — $\aleph_\alpha$ being regular — $A_{\mu_0} - \bigcup_{B_{\mu'} \in \mathfrak{B}'} B_{\mu'}$ is of power $\aleph_\alpha$ for every $\mathfrak{B}' \subseteq \mathfrak{S}$ $(\overline{\overline{\mathfrak{B}'}} < \aleph_\alpha)$.

(b) Let $\mathfrak{B}' \subseteq \mathfrak{B}$, $\overline{\overline{\mathfrak{B}'}} = \aleph_\alpha$. Put $A_{\mu_0} = \bigcup_{B_{\mu'} \in \mathfrak{B}'} (B_{\mu'} \cap S')$. Then $A_{\mu_0} \in \mathfrak{A}^*$, since if $\mathfrak{B}'' \subseteq \mathfrak{S}$ $(\overline{\overline{\mathfrak{B}''}} < \aleph_\alpha)$ is arbitrary, then there exists a μ_1 such that $B_{\mu_1} \in \mathfrak{B}'$ and $B_{\mu_1} \notin \mathfrak{B}''$, and so by (9)

$$\overline{\overline{(B_{\mu_1} \cap S') \cap B_{\mu''}}} < \aleph_\alpha \quad \text{for every} \quad B_{\mu''} \in \mathfrak{B}''.$$

Therefore, by the regularity of $\aleph_\alpha$, the set $(B_{\mu_1} \cap S') - \bigcup_{B_{\mu''} \in \mathfrak{B}''} B_{\mu''}$ is of power $\aleph_\alpha$. Consequently, the set $A_{\mu_0} - \bigcup_{B_{\mu''} \in \mathfrak{B}''} B_{\mu''}$ is non-empty.

Now we prove:

(13) If $A_\mu \in \mathfrak{A}^*$ and $\mu < \nu$, then $A_\mu \cap B_\nu$ is non-empty.

In fact, if $A_\mu \in \mathfrak{A}^*$, then $A_\mu \in \mathfrak{A}_\nu^*$ and therefore by (5) there is a $\varrho < \omega_\alpha$ such that $A_\mu = A_\varrho^\nu$ and x_ϱ^ν is a common element of A_μ and B_ν by (6) and (7).

(14) Suppose $\mathfrak{S}'\subseteq\mathfrak{S}$, $\overline{\overline{\mathfrak{S}'}}=\aleph_{\alpha+1}$. We have to prove that $S'=S-\bigcup_{B_\mu\in\mathfrak{S}'} B_\mu$ is of power less than $\aleph_{\alpha+1}$.

For if not, then $\overline{\overline{S'}}=\aleph_{\alpha+1}$ and by (12) there is a $\mu_0<\omega_{\alpha+1}$ such that $A_{\mu_0}\in\mathfrak{A}^*$ and $A_{\mu_0}\subseteq S'$. Since $\overline{\overline{\mathfrak{S}'}}=\aleph_{\alpha+1}$, there is a $\nu_0>\mu_0$ such that $B_{\nu_0}\in\mathfrak{S}'$. But then by (13) $A_{\mu_0}\cap B_{\nu_0}\neq 0$ what is a contradiction.

The system $\mathfrak{S}$ satisfies the requirements of Theorem 2 by (8), (9), (10) and (14). Q. e. d.

Theorem 2 implies Theorem 1 as follows:

Let S be a set, $\overline{\overline{S}}=\aleph_{\alpha+1}$ and suppose $\aleph_\alpha$ is regular. Then by Theorem 2 there exists a system $\mathfrak{S}$ satisfying the formulas (+)—(++++). Let $S=\{x_\nu\}_{\nu<\omega_{\alpha+1}}$ and $\mathfrak{S}=\{B_\nu\}_{\nu<\omega_{\alpha+1}}$ be well-orderings of type $\omega_{\alpha+1}$ of the sets S and $\mathfrak{S}$, respectively. We define $f(x_\nu)$ by induction on ν as follows:

Put $f(x_0)=0$ if $x_0\in B_0$ and $f(x_0)=B_0$ if $x_0\notin B_0$. Suppose that $f(x_\mu)$ is already defined for every $\mu<\nu$. Then there exists a least ϱ_0 such that $B_{\varrho_0}\neq f(x_\mu)$ for every $\mu<\nu$.

Put $f(x_\nu)=0$ if $x_\nu\in B_{\varrho_0}$ and $f(x_\nu)=B_{\varrho_0}$ if $x_\nu\notin B_{\varrho_0}$. Let S^* be the subset of those x_ν's for which $f(x_\nu)\neq 0$. Obviously $\overline{\overline{S^*}}=\aleph_{\alpha+1}$. Put $f^*(x_\nu)=$ $=f(x_\nu)\cap S^*$ for $x_\nu\in S^*$. The set-mapping $f^*(x)$ defined on S^* satisfies the requirements of Theorem 1. (++) and (+++) imply immediately that (o) and (oo) hold. Let now $S'\subseteq S^*$, $\overline{\overline{S'}}=\aleph_{\alpha+1}$. Put $\mathfrak{S}'=\{f(x_\nu)\}_{x_\nu\in\mathfrak{S}'}$. $\mathfrak{S}'$ is a subset of power $\aleph_{\alpha+1}$ of $\mathfrak{S}$. It follows from (++++) that the set $S'\cap\bigcup_{x_\nu\in\mathfrak{S}'} f(x_\nu)=$ $=S'\cap\bigcup_{x_\mu\in\mathfrak{S}'} f^*(x_\nu)$ is non-empty, hence S' is not free. So we have defined a set-mapping on S^* satisfying the requirements of Theorem 1. It results that one can define such a set-mapping on every set of power $\aleph_{\alpha+1}$, and the theorem is proved.

The general problem for set-mappings is the following: Let S be a set, $\overline{\overline{S}}=m$, and let $f(x)$ be a set-mapping defined on S, $\overline{\overline{f(x)}}<p$ for every $x\in S$.

Suppose that $\overline{\overline{\bigcap_{x\in S'} f(x)}}<q$ for every $S'\subseteq S$ $(\overline{\overline{S'}}\leqq r)$. Does then there exist a free subset of power s?

For the sake of brevity we now introduce the symbol $[[m, p, q, r]]\to s$ to indicate that this statement is true and, as usual, $[[m, p, q, r]]\not\to s$ indicates the negation of it.

Theorem 1 can be formulated with the help of this symbol as follows:

(*) $[[\aleph_{\alpha+1}, \aleph_{\alpha+1}, \aleph_\alpha, 2]]\not\to\aleph_{\alpha+1}$

if $\aleph_\alpha$ is regular. We wish to remark that we have some results for the gen-

eral case with P. ERDŐS and R. RADO which will be published later.[5] Here we want only to state the simplest unsolved problems for the case $m=p$, $r=2$.

(*) Theorem 7 of [2] states that $[[m, m, n, 2]] \to m$ if m is singular and $n<m$.

We mention that with the help of the so-called measure hypothesis[6] one can prove the same if m is inaccessible (*).

Thus we may suppose $m=\aleph_{\alpha+1}$. If $\aleph_\alpha$ is regular, then

(*) $[[\aleph_{\alpha+1}, \aleph_{\alpha+1}, \aleph_\alpha, 2]] \to \aleph_\alpha$,

and so Theorem 1 completes the discussion for this case. But if $\aleph_\alpha$ is singular, e. g. $\aleph_\alpha = \aleph_\omega$, we know only

(*) $[[\aleph_{\omega+1}, \aleph_{\omega+1}, \aleph_\omega, 2]] \to \aleph_0$

and the following problem remains open:

PROBLEM 1. $[[\aleph_{\omega+1}, \aleph_{\omega+1}, \aleph_\omega, 2]] \to \aleph_1$?

We mention here that similarly to the case of the set-mappings a general problem can be formulated, corresponding to the problem treated in Theorem 2, and there are some interesting cases when the two corresponding problems are not equivalent.[7]

We want to state another problem for set-mappings which we can not solve and therefore we formulate it only in the simplest case:

PROBLEM 2. Let S be a well-ordered set, $\overline{\overline{S}} = \aleph_1$, $\overline{S} = \omega_1$, and let $f(x)$ be a set-mapping defined on S, $\overline{\overline{f(x) \cap f(y)}} < \aleph_0$ for every $x, y \in S$ $(x \neq y)$. Does there then exist a free subset of type α for an arbitrary $\alpha < \omega_1$?

REMARK. We may suppose that $\overline{\overline{f(x)}} = \aleph_0$ for every $x \in S$. Theorem 8 of [2] assures the existence of a free subset of type ω and our Theorem 1 shows that the existence of a free subset of power ω_1 can not be proved.

Using the methods of the proof of Theorem 8 proved in Section 3 of this paper we can prove the weaker

THEOREM 3. *Under the conditions of Problem 2 there exists a free subset of type $\omega^2 + \omega \cdot n$ for every finite n.*

We omit the proof of this theorem.

Let now S be the set of the real numbers, and let $f(x)$ be a set-mapping defined on S such that $\overline{\overline{f(x)}} = \aleph_0$, $\overline{\overline{f(x) \cap f(y)}} < \aleph_0$ for every $x, y \in S$ $(x \neq y)$.

[5] See the forthcoming paper of P. ERDŐS, R. RADO and A. HAJNAL.

[6] See e. g. [4].

[7] All these results will be published in the paper mentioned above.

We have the following

THEOREM 4. *There exists a free subset S' of S which is everywhere dense.*

Theorem 4 assures the existence of a free set of every denumerable order type.

Theorem 4 is to be proved similarly to Theorem 6 of Section 3, and therefore here we omit the proof. See the remark after the proof of Theorem 6.

3. Results on the partition symbol. We repeat here the definition of the partition symbol in the special case needed for our purposes.

Let $\Theta, \Theta_1, \Theta_2, \ldots$ denote order types. The symbol $\Theta \to (\Theta_1, \Theta_2)^2$ indicates the following statement:

(′) Let S be an arbitrary ordered set of type Θ, and let $[S]^2 = J_1 \cup J_2$ be an arbitrary splitting of $[S]^2$. Then either there exists a subset $S_1 \subseteq S$ $(\bar{S}_1 = \Theta_1)$ such that $[S_1]^2 \subseteq J_1$, or there exists a subset $S_2 \subseteq S$ $(\bar{S}_2 = \Theta_2)$ such that $[S_2]^2 \subseteq J_2$. $\Theta \not\to (\Theta_1, \Theta_2)^2$ indicates the negation of this statement. Moreover, we are going to use the signs $\wedge$ (and) and $\vee$ (or) in the symbols in the following natural way: e. g. $\Theta \to (\Theta_1, \Theta_2 \vee \Theta_3)^2$ indicates the statement that under the conditions of (′) either there exists an S_1 $(S_1 \subseteq S)$ such that $[S_1]^2 \subseteq J_1$, or either there exists an S_2 $(S_2 \subseteq S, \bar{S}_2 = \Theta_2)$ such that $[S_2]^2 \subseteq J_2$, or there exists an S_3 $(S_3 \subseteq S, \bar{S}_3 = \Theta_3)$ such that $[S_3]^2 \subseteq J_2$.

Throughout this section λ and η denote the order types of the set of all real and all rational numbers under order by magnitude, respectively.

First of all we prove the following negative result:

THEOREM 5. $\omega_1 \not\to (\omega + 2, \omega_1)^2$. [8]

PROOF. Theorem 5 is an immediate consequence of Theorem 1. Let S be a well-ordered set, $\bar{S} = \omega_1$, and let $f(x)$ be a set-mapping defined on S satisfying the formulas (o), (oo), (ooo) of Theorem 1 for $\alpha = 0$. We may suppose by (o) that $y \in f(x)$ implies $y < x$.

We define J_1 and J_2 as follows: The pair $\{y, x\}$ $(y < x)$ belongs to J_1 if $y \in f(x)$ and it belongs to J_2 if $y \notin f(x)$. Then $\bar{S}_1 < \omega + 2$ by (oo) if $[S_1]^2 \subseteq J_1$ and $\bar{S}_2 < \omega_1$ by (ooo) if $[S_2]^2 \subseteq J_2$.

We are going to use the sign $\Theta_1 \leqq \Theta_2$ in the same sense as it is introduced in [1], p. 428.

[8] Theorem 5 would be compared with a result of [1] which states that (*) $\omega_{\alpha+1} \to (\omega_\alpha + 1, \omega_{\alpha+1})^2$ if $\aleph_\alpha$ is regular. This is a corollary of Theorem 39 of [1].

It is obvious that we can prove similarly to Theorem 5 that (*) $\omega_{\alpha+1} \not\to (\omega_\alpha + 2, \omega_{\alpha+1})^2$ if $\aleph_\alpha$ is regular, but here we do not want to discuss the case of greater cardinals because we can prove almost all the other results only in the special cases for λ and ω_1.

('') Let S be an ordered set, $A_1, A_2 \subseteq S$. We say that $A_1 \prec A_2$ if $x \in A_1, y \in A_2$ implies that $x < y$ in the ordering of S.

The type Φ is briefly called a real type if $\overline{\overline{\Phi}} > \aleph_0$ and $\omega_1, \omega_1^* \not\leqq \Phi$.

It is obvious that λ and every type $\lambda_1 \leqq \lambda$ $(\overline{\overline{\lambda}}_1 > \aleph_0)$ are real types.

We need some preliminary lemmas.

LEMMA 1. *Let S be an ordered set of the real type Φ. There exists a sequence $J = \{A_n\}_{n<\omega}$ of subsets of S such that*

(x) $\overline{\overline{A}}_n > \aleph_0$ *for every* $n < \omega$,

(xx) *the set J is ordered by the relation $\prec$ introduced in ('') and is of type η.*

LEMMA 2. *Let S be the set of the real numbers, and let $I_0, \ldots, I_n, \ldots$ be the sequence of all intervals with rational endpoints. There exists a sequence $\{A_n\}_{n<\omega}$ of subsets of S such that*

$(\cdot)$ $\overline{\overline{A}}_n > \aleph_0$ *for every* $n < \omega$,

$(\cdot\cdot)$ $A_n \cap A_m = 0$ *if* $m \neq n$,

$(\cdot\cdot\cdot)$ $A_n \subseteq I_n$ *for every* $n < \omega$.

Both lemmas are well known and we omit the proofs.

LEMMA 3. *Let S be a set, $\overline{\overline{S}} > \aleph_0$, and let $\{A_n\}_{n<\omega}$ be a sequence of type ω of subsets of S such that $\overline{\overline{A}}_n > \aleph_0$ for every $n < \omega$ and $A_n \cap A_m = 0$ if $n \neq m$. Let further $[S]^2 = J_1 \cup J_2$ be an arbitrary splitting of $[S]^2$. Then either*

(α) *there exists a subset S' of S such that $[S']^2 \subseteq J_2$ and $S' \cap A_n \neq 0$ for every $n < \omega$, or*

($\alpha\alpha$) *there exist two subsets S_1 and S_2 of S such that $S_1 \cap S_2 = 0$, $\overline{\overline{S}}_1 > \aleph_0$, $\overline{\overline{S}}_2 > \aleph_0$, $S_i \subseteq A_{n_i}$ for $i = 1, 2$ and for suitable $n_1 \neq n_2$, and for every $x \in S_1$ the set $\{y : y \in S_2, \{x, y\} \in J_2\}$ is of power at most $\aleph_0$.*

PROOF. Let $R_i(x, S')$ denote the set $\{y : y \in S', \{x, y\} \in J_i\}$ for $i = 1, 2$, $x \in S$ and $S' \subseteq S$. Then either

(β) there exists an element $x_0 \in A_0$ such that $\overline{\overline{R_2(x_0, A_n)}} > \aleph_0$ for every $0 < n < \omega$, or

($\beta\beta$) corresponding to every element x of A_0 there exists an index $0 < n(x) < \omega$ such that $\overline{\overline{R_2(x, A_{n(x)})}} \leqq \aleph_0$. If ($\beta\beta$) holds, then we split A_0 into the sum of $\aleph_0$ classes B_n. We put $x \in B_n$ if $n(x) = n$. Obviously $A_0 = \bigcup_n B_n$, hence there is an index $0 < n_0 < \omega$ such that $\overline{\overline{B}}_{n_0} > \aleph_0$. Then $S_1 = B_{n_0}$, $S_2 = A_{n_0}$ satisfy the requirements of ($\alpha\alpha$).

Thus we may suppose that (β) holds. Put now $A_0^1 = R_2(x_0, A_1), \ldots, A_n^1 = R_2(x_0, A_{n+1}), \ldots$.

Repeating the above consideration for this new sequence we may suppose again that (β) holds, and so we obtain an element $x_1 \in A_1$ and a new sequence $A_0^2, \ldots, A_n^2, \ldots$ and so on. Finally, we get a sequence $\{x_n\}_{n<\omega}$ such that $x_n \in A_n$ and $\{x_n, x_m\} \in J_2$ for every n, m $(n \neq m)$, and then $S' = \{x_n\}_{n<\omega}$ satisfies (α).

Now we are going to prove

THEOREM 6. *$\Phi \rightarrow (\omega \cdot n, \alpha)^2$ for an arbitrary real type Φ where n is finite and $\alpha < \omega_1$.*

REMARK. This theorem is a generalization of Theorem 31, (28)—(29) of [1] ($\Phi \rightarrow (\omega + n, \omega \cdot m)^2$ for finite n and m, and $\Phi \rightarrow (\omega, \alpha)^2$ for every $\alpha < \omega_1$) and gives a solution of the problem $\lambda \rightarrow (\omega \cdot 2, \omega \cdot 2)^2$ stated in the introduction of [1], p. 428. However, it is not a solution of the general problem. The simplest unsolved problem here is

PROBLEM 3. $\lambda \rightarrow (\omega^2, \omega^2)^2$?

PROOF OF THEOREM 6. We prove the theorem by induction on n. For $n = 1$ the theorem is known (see [1], Theorem 31, (29)). Suppose that it holds for an $n \geqq 1$ for every real type Φ and for every ordinal number $\alpha < \omega_1$. Let S be an ordered set. Let $\bar{S} = \Phi$ be a real type. Let $\{A_n\}_{n<\omega}$ be a sequence of subsets of S satisfying the requirements (x) and (xx) of Lemma 1 for α instead of η.[9] Let now $[S]^2 = J_1 \cup J_2$ be an arbitrary splitting of $[S]^2$.

Then the conditions of Lemma 3 are satisfied, and therefore either (α) or ($\alpha\alpha$) holds.

If (α) holds, then we may supose that $\overline{\overline{S' \cap A_n}} = 1$ for every $n < \omega$, and then we have $S' \subseteq S$, $[S']^2 \subseteq J_2$, $\bar{S'} = \alpha$ by the formula (xx) of Lemma 1.

Thus if (α) holds, the theorem is true, so we may suppose that ($\alpha\alpha$) holds.

If ($\alpha\alpha$) holds, then we have two disjoint subsets S_1 and S_2 of S such that $\bar{\bar{S}}_1 > \aleph_0$, $\bar{\bar{S}}_2 > \aleph_0$ and $\overline{\overline{R_2(x, S_2)}} \leqq \aleph_0$ for every $x \in S_1$. Since $S_1 \subseteq A_{n_i}$, we may suppose that either $S_1 \prec S_2$ or $S_2 \prec S_1$.

Put $\Phi_1 = \bar{S}_1$. Obviously Φ_1 is a real type, and therefore we have

$$\Phi_1 \rightarrow (\omega \cdot n, \alpha)^2.$$

Thus, if there is no $S' \subseteq S$ $(\bar{S'} = \alpha, [S'] \subseteq J_2)$, then there exists an $S_1^* \subseteq S_1$ $(\bar{S}_1^* = \omega \cdot n)$ such that $[S_1^*]^2 \subseteq J_1$.

Put $R = \bigcup_{x \in S_1^*} R_2(x, S_2)$. Then $\bar{\bar{R}} \leqq \aleph_0$, and therefore $\overline{\overline{S_2 - R}} > \aleph_0$. Put

[9] It is obvious that if Lemma 1 holds for η, then it holds for an arbitrary denumerable order type.

$\Phi \rightarrow (\omega \cdot n, \eta)^2$ is not true. The induction proof breaks down at the first step, since, as it is well known, $\lambda \nrightarrow (\omega, \eta)^2$.

$S_2' = S_2 - R$, $\Phi_2 = \bar{S}_2'$. Φ_2 is a real type, and so we have $\Phi_2 \to (\omega, \alpha)^2$. Thus we may suppose that there exists an $S_2^* \subseteq S_2'$ $(\bar{S}_2^* = \omega)$ such that $[S_2^*]^2 \subseteq J_1$.

Put $S^* = S_1^* \cup S_2^*$. Then $S^* \subseteq S$, $[S^*]^2 \subseteq J_1$, since $[S_1^*]^2, [S_2^*]^2 \subseteq J_1$ and $\{xy\} \in J_1$ if $x \in S_1^*, y \in S_2^*$. The type of S^* is either $\omega \cdot n + \omega$ or $\omega + \omega \cdot n$, i. e. $\bar{S}^* = \omega \cdot (n+1)$. Q. e. d.

Remark. Theorem 4 of Section 2 can be proved similarly. If we use Lemma 2 instead of Lemma 1, then applying Lemma 3 we get either a free set everywhere dense if (α) holds or a contradiction if $(\alpha\alpha)$ holds.

One may conjecture that $\lambda \to (\alpha, \alpha)^2$ holds for every $\alpha < \omega_1$. The theorem just proved gives a weaker result in this direction and seems very likely not to be best-possible.

Let S be an ordered set, and let $[S]^2 = J_1 \cup J_2$ be a splitting of $[S]^2$. If there exists an $S' \subseteq S$ $([S']^2 \subseteq J_i,\ \bar{S}' = \Theta)$, we shall briefly say that J_i contains a complete set of type Θ. The conjecture mentioned above can be stated as follows: If $\bar{S} = \lambda$ and the splitting is arbitrary, then for every $\alpha < \omega_1$ one of the classes J_i contains a complete set of type α, and one of the classes J_i contains a complete set of type α^*. (Since the theorem is true for $\lambda^* = \lambda$ too.) It was not known whether for an arbitrary α there exists a J_i which contains a complete set of type α either increasing or decreasing, i. e. a complete set either of type α or of type α^*. Now we can prove a more general theorem which assures this and is best-possible of its kind, however, it does not imply the general conjecture. This is

Theorem 7. $\Phi \to (\alpha \vee \alpha^*, \eta)^2$ *for every real type* Φ *and for every* $\alpha < \omega_1$.

Theorem 7 can be proved essentially with the same idea as Theorem 6. However, there are some technical difficulties, and first we need some lemmas concerning denumerable order types.

Let $\Theta_0, \dots, \Theta_n, \dots$ be a sequence of order types, $\Theta_0 + \dots + \Theta_n + \dots$ and $\dots + \Theta_n + \dots + \Theta_0$ are briefly called the ω-sum and the ω^*-sum of the sequence, respectively.

Now we define an increasing sequence O_ϱ of sets of denumerable order types by induction on ϱ for every $\varrho < \omega_1$ as follows. Put $O_0 = \{\omega, \omega^*\}$. Suppose that O_β is defined for every $\beta < \varrho$. Put $O_\varrho^+ = \bigcup_{\beta < \varrho} O_\beta$. Let O_ϱ consist of the elements of O_ϱ^+ and of the ω-sums and ω^*-sums of all subsequences of O_ϱ^+. Put finally $O = \bigcup_{\varrho < \omega_1} O_\varrho$.

For an arbitrary $\Theta \in O$ we define the order of Θ as the least ordinal number ϱ for which $\Theta \in O_\varrho$. Let $\varrho(\Theta)$ denote the order of ϱ.

We need the following lemmas:

(I) Let $\{\Theta_n\}_{n<\omega}$ be a sequence of denumerable order types and suppose

$\Theta_n \in O$, $\varrho(\Theta_n) \geqq \varrho_n$ for $n=1,2,\ldots$. Then the order of the ω-sum as well as the order of the ω^*-sum of the sequence is greater than ϱ_n for every n.

(II) If α is an arbitrary ordinal number, $\alpha<\omega_1$, then there exists a ϱ_0 (depending on α) such that if $\varrho(\Theta)>\varrho_0$ for a $\Theta \in O$, then either $\alpha \leqq \Theta$ or $\alpha^* \leqq \Theta$.

The verification of (I) and (II) is very simple, and so we may omit the proofs.[10]

Now we need the following

LEMMA 4. *Let S be an ordered set of real type Φ, and let $[S]^2 = J_1 \cup J_2$ be an arbitrary splitting of $[S]^2$. Then one of the following statements holds:*

(γ) *There exists an $S' \subseteq S$, $[S']^2 \subseteq J_2$, $\bar{S}' = \eta$.*

($\gamma\gamma$) *There exists a sequence $\{T_k\}_{k<\omega}$ of subsets of S satisfying the following conditions:*

(v) $T_0 \supseteq \cdots \supseteq T_k \supseteq \cdots$.

Put $U_k = T_k - T_{k+1}$.

(vv) $\bar{\bar{U}}_k > \aleph_0$ *for every* $k<\omega$.

(vvv) *The set $\{U_k\}_{k<\omega}$ is ordered by the relation $\prec$ introduced in* (″).

(vvvv) $\overline{\overline{R_2(x, T_{k+1})}} \leqq \aleph_0$ *if $x \in U_k$ for every $k<\omega$.*[11]

PROOF. Now we suppose that (γ) is false and we define two sequences $\{U_k\}_{k<\omega}$, $\{V_k\}_{k<\omega}$ by induction on k as follows: Let $\{A_n^0\}_{n<\omega}$ be a sequence satisfying the requirements of Lemma 1. We apply Lemma 3 to this sequence; it is obvious that if (α) holds, then (γ) is true. So we may suppose that ($\alpha\alpha$) holds; let S_1^0, S_2^0 be the sets satisfying ($\alpha\alpha$). Put

(1′) $$U_0 = S_1^0, \quad V_0 = S_2^0.$$

(2′) $\bar{\bar{S}}_2^0 > \aleph_0$, and therefore $\Phi_1 = \bar{V}_0$ is a real type.

Suppose that $U_0, \ldots, U_k$, $V_0, \ldots, V_k$ are already defined in such a way that $\Phi_{k+1} = \bar{V}_k$ is a real type.

Let $\{A_n^{k+1}\}_{n<\omega}$ be a sequence satisfying the requirements of Lemma 1 with V_k and Φ_k instead of S and Φ, respectively.

We apply again Lemma 3 to this sequence and similarly as in the first step we may suppose that ($\alpha\alpha$) holds. Let S_1^{k+1}, S_2^{k+1} be the sets satisfying ($\alpha\alpha$) in this case and put

(3′) $$U_{k+1} = S_1^{k+1}, \quad V_{k+1} = S_2^{k+1}, \quad \Phi_{k+1} = \bar{V}_{k+1}.$$

[10] The author does not know whether the above-mentioned results are known. It is to be remarked that one can prove that every denumerable order type Θ belongs to O, provided $\eta \not\leqq \Theta$. See a forthcoming paper of P. ERDŐS and A. HAJNAL.

[11] $R_2(x, S')$ is defined in the proof of Lemma 3.

We have

(4′) $\overline{\overline{V}}_{k+1}=\overline{\overline{S}}_2^{k+1}>\aleph_0$, and so Φ_{k+2} is a real type.

Thus U_k, V_k are defined for every $k<\omega$. Put

(5′) $$T_k=\bigcup_{k'\geqq k} U_{k'}.$$

It is obvious that (v) holds. We have

(6′) $\overline{\overline{U}}_k=\overline{\overline{S}}_1^k>\aleph_0$ for every $k<\omega$ by (1′) and (3′′′), i. e. (vv) holds.

The sets $\{A_n^k\}_{n\prec\omega}$ are ordered with respect to the relation $\prec$ by Lemma 1 for every $k<\omega$. By Lemma 3 $S_1^k\subseteq A_{n_1}^k$, $S_2^k\subseteq A_{n_2}$, and so by (1′) and (3′) we have

(7′) either $U_k\prec V_k$ or $V_k\prec U_k$.

On the other hand, it follows from (1′) and (3′) that

(8′) $U_{k'}\subseteq V_k$ for $k'>k$.

Comparing (7′) and (8′) we get

(9′) either $U_k\prec U_{k'}$ or $U_{k'}\prec U_k$ for every $k\neq k'$, i. e. (vvv) holds.

From (5′) and (8′) we get

(10′) $T_{k+1}\subseteq V_k$ for every $k<\omega$.

Hence $R_2(x, T_{k+1})\subseteq R_2(x, V_k)$ for every $k<\omega$.
It follows from (1′) and (3′) that

(11′) $\overline{\overline{R_2(x, T_{k+1})}}\leqq\aleph_0$ if $x\in U_k$ for every $k<\omega$, i. e. (vvvv) holds.

PROOF OF THEOREM 7. By (II) it is enough to prove the following statement:

Let S be an ordered set of real type Φ, and let $[S]^2=J_1\cup J_2$ be an arbitrary splitting of $[S]^2$. Then one of the following statements holds:

(δ) There is an $S'\subseteq S$, $[S']^2\subseteq J_2$, $\overline{S'}=\eta$.

($\delta\delta$) For every $\varrho<\omega_1$ there exists a $\Theta\in O$ and an $S''\subseteq S$ such that $[S'']^2\subset J_1$, $\overline{S''}=\Theta$ and $\varrho(\Theta)\geqq\varrho$.

Suppose that (δ) is false. We have to prove that then ($\delta\delta$) holds: We prove this by induction on ϱ. For $\varrho=0$ the theorem is well known.[12]

[12] For $\varrho=0$ the theorem states that $\Phi\to(\omega\vee\omega^*, \eta)^2$, but this is the same as $\Phi\to(\aleph_0, \eta)^2$, but we have $\aleph_1\to(\aleph_0, \aleph_1)^2$ (see. e. g. [1], Theorem 3, (i)), and every set of real type has a subset of type η.

Let $\varrho > 0$, and suppose that the theorem is true for every $\varrho' < \varrho$. Let $\varrho_0, \ldots, \varrho_n, \ldots$ be a sequence of ordinal numbers less than ϱ such that every ordinal number which is greater than every ϱ_n is not less than ϱ. (Put $\varrho_n = \varrho - 1$ if ϱ is of the first kind, and let ϱ_n be a sequence $\varrho_n \to \varrho$ if ϱ is of the second kind.)

Now we can apply Lemma 4. If (γ) holds, then (δ) is true in contrary to our assumption. So we may suppose that ($\gamma\gamma$) holds. Let U_k, T_k be the sets satisfying the requirements of ($\gamma\gamma$).

(12′) The set $\{U_k\}_{k<\omega}$ is ordered with respect to the relation $\prec$ by (vvv), and so there exists a subsequence $\{U_{k_n}\}_{n<\omega}$ which is either increasing or decreasing.

Now we define a sequence $\{S_n\}_{n<\omega}$ of subsets of S as follows:

The set U_{k_0} has a real type by (vv), and so by the induction hypothesis there exists an S_0 such that

(13′) $S_0 \subseteq U_{k_0}$, $[S_0]^2 \subseteq J_1$, $\bar{S}_0 = \Theta_0$ where $\Theta_0 \in O$ and $\varrho(\Theta_0) \geqq \varrho_0$.

Suppose that the sets S_i ($i < n$) are already defined in such a way that $S_i \subseteq U_{k_i}$ and $\bar{S}_i = \Theta_i$, $\Theta_i \in O$ for every $i < n$.

Then the set $U_{k_n} - \bigcup_{i=0}^{n-1} \bigcup_{x \in S_i} R_2(x, T_{k_i+1})$ is of power $> \aleph_0$ by (vv) and (vvvv), this means that it has a real type and by the induction hypothesis there exists an S_n such that

(14′) $S_n \subseteq U_{k_n} - \bigcup_{i=0}^{n-1} \bigcup_{x \in S_i} R_2(x, T_{k_i+1})$, $[S_n]^2 \subseteq J_1$, $\bar{S}_n = \Theta_n$ for a $\Theta_n \in O$ and $\varrho(\Theta_n) \geqq \varrho_n$.

Now S_n is defined for every n (since $S_n \subseteq U_{k_n}$, $\Theta_n \in O$ holds) and it is proved that S_n satisfies (14′) for every $n < \omega$.

Put $S'' = \bigcup_{n<\omega} S_n$. It follows immediately from the first two formulas of (14′) that

(15′) $$[S'']^2 \subseteq J_1.$$

Put

(16′) $$\Theta = \bar{S}'',$$

by (12′) Θ is either the ω-sum or the ω^*-sum of the sequence $\{\Theta_n\}_{n<\omega}$, and therefore from the lemma (I) we get

(17′) $\varrho(\Theta) > \varrho_n$ for every n, hence $\varrho(\Theta) \geqq \varrho$ by the definition of the sequence ϱ_n.

(15′), (16′) and (17′) prove Theorem 7.

In [1], Theorem 33, it has been proved that $\omega_1 \to (\alpha, \alpha)^2$ holds for $\alpha < \omega \cdot 2$. One can also conjecture that $\omega_1 \to (\alpha, \alpha)^2$ holds for every $\alpha < \omega_1$.

Here we can not prove as much as in the case of real types, but perhaps the following result is of interest:

THEOREM 8. $\omega_1 \to (\omega \cdot n, \omega \cdot 2)^2$ *for every finite n.*

The simplest unsolved problems here are the following:

PROBLEM 4. $\omega_1 \to (\omega^2, \omega \cdot 2)^2$? $\omega_1 \to (\omega \cdot 3, \omega \cdot 3)^2$?[13]

For the proof of Theorem 8 some preliminary results are needed.

If $\{S_\alpha\}_{\alpha<\varphi}$ is an arbitrary sequence of sets, let $\bigotimes_{\alpha<\varphi} S_\alpha$ denote the set of those pairs $\{x, y\}$ for which there exists a pair of ordinal numbers $\alpha \neq \beta$ such that $x \in S_\alpha$, $y \in S_\beta$. (If φ is finite, we use the notations $S_1 \otimes S_2$, $S_1 \otimes S_2 \otimes S_3, \ldots$.)

First we need the following lemmas:

(III) Let S be an ordered set, $\bar{S} = \omega_1$, and let $[S]^2 = J_1 \cup J_2$ be an arbitrary splitting of $[S]^2$. Then one of the following statements holds:

(ε) There exists for every $n < \omega$ a subset $S_n \subseteq S$ such that $\bar{S}_n = \omega \cdot n$, $[S_n]^2 \subseteq J_1$.

($\varepsilon\varepsilon$) There exist subsets $P, R \subseteq S$ such that $\bar{\bar{P}} = \aleph_0$, $\bar{\bar{R}} > \aleph_0$ and $P \otimes R \subseteq J_2$.

Suppose that ($\varepsilon\varepsilon$) is false. We prove (ε) by induction on n. For $n = 1$ the statement is a corollary of $\aleph_1 \to (\aleph_0, \aleph_1)^2$. Suppose that it is true for an $n \geqq 1$. We have to prove that there is a set S_{n+1} satisfying (ε) (with $n+1$ instead of n).

First of all we have to verify some simple statements ((18′), (19′) and (20′)) not depending on the induction hypothesis.

(18′) Suppose $Q_0, Q_1 \subseteq S$, $\bar{\bar{Q}}_0 = \aleph_0$, $\bar{\bar{Q}}_1 > \aleph_0$. Then either ($\varepsilon\varepsilon$) holds or there exists an $x_0 \in Q_0$ such that $\overline{\overline{R_1(x_0, Q_1)}} > \aleph_0$.

For if $\overline{\overline{R_1(x, Q_1)}} \leqq \aleph_0$ for every $x \in Q_0$, then the sets $P = Q_0$, $R = Q_1 - \bigcup_{x \in Q_0} R_1(x, Q_1)$ satisfy ($\varepsilon\varepsilon$).

(19′) Suppose $Q_l \subseteq S$ for every $l < \omega$, $Q \subseteq S$, $Q_l \cap Q_{l'} = 0$ for $l \neq l'$, $l, l' < \omega$, and $\bar{\bar{Q}}_l = \aleph_0$ for every $l < \omega$, $\bar{\bar{Q}} > \aleph_0$. Then either ($\varepsilon\varepsilon$) holds for a $P \subseteq Q_l$, $R \subseteq Q$, or there exists an element x_0 of Q such that $\overline{\overline{R_1(x_0, Q_l)}} = \aleph_0$ for every $l < \omega$.

For if for every $x \in Q$ there exists a subscript $l(x)$ such that $\overline{\overline{R_1(x, Q_{l(x)})}} < \aleph_0$, then there exists a subset Q' of Q and an $l_0 < \omega$ such that $Q' = \{x : x \in Q$,

[13] It is possible that the proof of $\omega_1 \to (\omega \cdot n, \omega \cdot 2 + m)^2$ does not need new methods, but the problems stated above seem to need new ideas.

$l(x)=l_0\}$ and $\overline{\overline{Q'}}>\aleph_0$. But then there exists a subset $Q''\subseteq Q'$ such that $R_1(x, Q_{l_0})=R_1(x', Q_{l_0})$ for every pair $x, x'\in Q''$ and $\overline{\overline{Q''}}>\aleph_0$.

Therefore in this case $P=Q_{l_0}-\bigcup_{x\in Q''} R_1(x, Q_{l_0})$ and $R=Q''$ satisfy $(\varepsilon\varepsilon)$.

(20′) Suppose $S'\subseteq S$, $\overline{\overline{S'}}>\aleph_0$. Then either there exists a subset $S^*\subseteq S'$, $\overline{\overline{S^*}}>\aleph_0$, $[S^*]^2\subseteq J_2$ (and as a corollary of this $(\varepsilon\varepsilon)$ holds), or there exists an $S^{**}\subseteq S'$ such that $\overline{\overline{R_1(x, S^{**})}}>\aleph_0$ for every $x\in S^{**}$.

Put $S''=\{x: \overline{\overline{R_1(x, S')}}\leqq\aleph_0,\ x\in S'\}$. If $\overline{\overline{S''}}\leqq\aleph_0$, then put $S^{**}=S'-S''$. If $\overline{\overline{S''}}>\aleph_0$, then it is easy to define by transfinite induction a subset S^* of it satisfying the requirements.[14]

Let S_n be a set satisfying (ε). Then there exists a sequence $S_n^1, \dots, S_n^n$ such that

(21′) $$\overline{\overline{S_n^l}}=\omega \quad\text{for}\quad l=1,\dots,n \quad\text{and}\quad S_n^1\prec\cdots\prec S_n^n,\quad S_n=\bigcup_{l=1}^{n} S_n^l.$$

Let further S' be a set such that

(22′) $$S'\subseteq S,\quad \overline{\overline{S'}}=\aleph_1,\quad S_n\prec S'.$$

Put

(23′) $$\{x_m^l\}_{m<\omega}=S_n^l \quad\text{for}\quad l=1,\dots,n \quad\text{and}\quad \{x_\alpha\}_{\alpha<\omega_1}=S'$$

where $x_m^l<x_{m'}^l$ for $m<m'$ and $x_\alpha<x_\beta$ for $\alpha<\beta$ in the given well-ordering of S.

Now we are going to define the subsequences $\{x_{m_k}^l\}_{k<\omega}\subseteq S_n^l$ for $l=1,\dots,n$, $\{x_{\alpha_k}\}_{k<\omega}\subseteq S'$ by induction on k as follows:

(24′) First we define the elements $x_{m_0}^l$ for $l=1,\dots,n$ by induction on l. Applying (18′) for $Q_0=S_n^1$, $Q_1=S'$, we get an element $x_{m_0}^1\in S_n^1$ such that $\overline{\overline{R_1(x_{m_0}^1, S')}}>\aleph_0$. Are the elements $x_{m_0}^1,\dots,x_{m_0}^{l-1}$ already defined in such a way that $\bigcap_{j=1}^{l-1} R_1(x_{m_0}^j, S')$ is of power $>\aleph_0$, we define $x_{m_0}^l\in S_n^l$ applying (18′) again to $Q_0=S_n^l$, $Q_1=\bigcap_{j=1}^{l-1} R_1(x_{m_0}^j, S')$ so that $\bigcap_{j=1}^{l} R_1(x_{m_0}^j, S')$ is of power $>\aleph_0$.

Put $U_0=\bigcap_{i=1}^{n} R_1(x_{m_0}^l, S')$. We have

(25′) $$x_{m_0}^l\in S_n^l \quad\text{for}\quad l=1,\dots,n \quad\text{and}\quad \overline{\overline{U_0}}>\aleph_0.$$

Now we apply (20′) to $U_0=S'$. It follows that there exists a V_0 such that

(26′) $$V_0\subseteq U_0,\quad \overline{\overline{V_0}}=\aleph_1 \quad\text{and}\quad \overline{\overline{R_1(x, V_0)}}=\aleph_1 \quad\text{for every}\quad x\in V_0.$$

[14] The statements proved in (18′), (19′), (20′) are well known, we only repeated the proofs for the convenience of the reader.

Put $T_0^l = \{x_m^l\}_{m_0 < m < \omega}$ for $l = 1, \dots, n$.

Applying (19') to $T_0^l = Q_l$, $V_0 = Q$ (for $l \leqq n$ instead of $l < \omega$), we get an x_{α_0} such that

(27') $$x_{\alpha_0} \in V_0, \quad \overline{\overline{R_1(x_{\alpha_0}, T_0^l)}} = \aleph_0 \quad \text{for} \quad l = 1, \dots, n.$$

Put $S'(1) = R_1(x_{\alpha_0}, V_0)$. We have

(28') $\overline{\overline{S'(1)}} = \aleph_1$ and $S'(1) \subseteq U_0$ by (26').

Put now $S_n^l(1) = R_1(x_{\alpha_0}, T_0^l)$ for $l = 1, \dots, n$.

From the formulas (24')—(28') it follows that the formulas (21'), (22'), (23') hold if we replace S', S_n^l $(l = 1, \dots, n)$ by $S'(1), S_n^l(1)$ $(l = 1, \dots, n)$, respectively. Thus we may continue this construction for $k = 1$ and so on. So we may suppose that the elements x_k^l, x_{α_k} as well as the sets U_k, V_k, T_k, $S'(k+1), S_n^l(k+1)$ are defined for every $k < \omega$ and $l = 1, \dots, n$ and that the formulas (24')—(28') hold for them with k instead of 0, respectively.

Put $S_{n+1}^l = \{x_{m_k}^l\}_{k<\omega}$ for $l = 1, \dots, n$, $S_{n+1}^{n+1} = \{x_{\alpha_k}\}_{k<\omega}$ and $S_{n+1} = \bigcup_{l=1}^{n+1} S_{n+1}^l$. Now we have

(29') $S_{n+1}^l \subseteq S_n^l$ for $l = 1, \dots, n$ by (25'),

(30') $S_{n+1}^{n+1} \subseteq S'$ by (24'), (26') and (27').

If follows from (21') and (22') that

(31') $$S_{n+1}^1 \prec \dots \prec S_{n+1}^{n+1},$$

and since we have $\overline{\overline{S}}_{n+1}^l = \aleph_0$, we get

(32') $$\overline{S}_{n+1} = \omega \cdot (n+1),$$

i. e. the second requirement of (ε) holds.

We have to prove that $[S_{n+1}]^2 \subseteq J_1$.

Put $S'_{n+1} = \bigcup_{l=1}^{n} S_{n+1}^l$ and $S''_{n+1} = S_{n+1}^{n+1}$. It follows from (29') that $S'_{n+1} \subseteq S_n$, hence we have

(33') $$[S'_{n+1}]^2 \subseteq J_1.$$

Suppose now that $x \in S'_{n+1}$, $y \in S''_{n+1}$. Then x has the form $x_{m_k}^l$ for suitable l and k, and y has the form $x_{\alpha_{k'}}$. We distinguish two cases:

(a) $k \leqq k'$, (b) $k > k'$.

If (a) holds, then $x_{\alpha_{k'}} \in U_k$ by (26'), (27') and (28'), and therefore $x_{\alpha_{k'}} \in R_1(x_{m_k}^l, S'(k))$ by (24'), i. e. $\{x_{m_k}^l, x_{\alpha_{k'}}\} \in J_1$.

If (b) holds, then $x^l_{m_k} \in S^l_n(k) \subseteq S^l_n(k')$ by (25′) and (28′), hence $\{x^l_{m_k}, x_{\alpha_{k'}}\} \in J_1$ holds in this case too.

Thus we have

(34′) $$S'_{n+1} \otimes S''_{n+1} \subseteq J_1.$$

Finally, if $x_{\alpha_k}, x_{\alpha_{k'}}$ $(k<k')$ are two arbitrary elements of S''_{n+1}; then $x_{\alpha_{k'}} \in S'(k') \subseteq S'(k+1)$ by (24′), (26′), (27′) and (28′). It follows from (27′) that $\{x_{\alpha_k}, x_{\alpha_{k'}}\} \in J_1$, i. e.

(35′) $$[S''_{n+1}]^2 \subseteq J_1.$$

Comparing (33′), (34′) and (35′) we get

(36′) $$[S_{n+1}]^2 \subseteq J_1,$$

hence the first requirement of (ε) holds and (III) is proved. As a consequence of (III) we get the following lemma:

(IV) Let S be a set. Under the conditions of lemma (III), either the statement (ε) of (III) holds, or

$(\varepsilon\varepsilon\varepsilon)$ there exists a sequence $\{P_l\}_{l<\omega}$ of subsets of S such that $\bar{P_l} = \aleph_0$ for every $l<\omega$, $P_0 \prec \cdots \prec P_l \prec \cdots$ and $\bigotimes_{l<\omega} P_l \subseteq J_2$.

Suppose that (ε) is false. We define the sequence $\{P_l\}_{l<\omega}$ by induction on l as follows: By lemma (III) there exist subsets P, R satisfying $(\varepsilon\varepsilon)$.

Put $P_0 = P$, $Q_0 = R - \bigcup_{x' \leqq x,\, x \in P_0} x'$.

Suppose that P_j and Q_j are already defined for $j \leqq l$ in such a way that $\bar{P_j} = \aleph_0$, $\bar{Q_j} = \aleph_1$. We apply (III) again to Q_l instead of S. Since (ε) is supposed to be false, we get sets P' and R' satisfying $(\varepsilon\varepsilon)$ (with Q_l instead of S).

Put $P_{l+1} = P$ and $Q_{l+1} = R - \bigcup_{x' \leqq x,\, x \in \bigcup_{j=1}^{l} P_j} x'$. Then $\bar{P}_{l+1} = \aleph_0$, $\bar{\bar{Q}}_{l+1} = \aleph_1$, and so the sequences $\{P\}_{l<\omega}$, $\{Q_l\}_{l<\omega}$ are defined.

It follows from the definition that $\bar{P_l} = \aleph_0$ for every $l<\omega$ and $P_l \prec P_{l'}$ if $l<l'<\omega$ and further $P_{l'} \subseteq Q_l$ for every $l'>l$ and for every $l<\omega$.

Therefore we have

$$\bigotimes_{l<\omega} P_l = \bigcup_{l<\omega} (P_l \otimes \bigcup_{l'>l} P_{l'}) \subseteq \bigcup_{l<\omega} P_l \otimes Q_l \subseteq J_2$$

and (IV) is proved.

Proof of Theorem 8. We prove the theorem by induction on n. For the case $n=1$ the theorem $\omega_1 \to (\omega, \omega\cdot 2)^2$ is a corollary of $\aleph_1 \to (\aleph_0, \aleph_1)^2$.[15] Suppose that the theorem is true for an $n \geqq 1$.

[15] See e. g. [1], Theorem (3), (I).

Let S be an ordered set, $\bar{S}=\omega_1$, and let $[S]^2=J_1\cup J_2$ be an arbitrary splitting of $[S]^2$. We have to prove that one of the following statements holds:

($\varkappa$) There is a subset $S_{n+1}\subseteq S$ such that $[S_{n+1}]^2\subseteq J_1$ and $\bar{S}_{n+1}=\omega\cdot(n+1)$.

($\varkappa\varkappa$) There is a subset $S'\subseteq S$ such that $[S']^2\subseteq J_2$ and $\bar{S}'=\omega\cdot 2$.

We may apply lemma (IV). If (ε) holds, then ($\varkappa$) is true. Thus we may suppose that there exists a sequence $\{P_l\}_{l<\omega}$ satisfying the requirements of ($\varepsilon\varepsilon\varepsilon$), i. e.

($\times$) $\bar{P}_l=\omega$ for $l<\omega$,

($\times\times$) $P_0\prec\cdots\prec P_l\prec\cdots$,

($\times\times\times$) $\bigotimes_{l<\omega} P_l\subseteq J_2$,

and we may suppose that

($\times\times\times\times$) $[P_l]^2\subseteq J_1$ for $l<\omega$.[16]

Suppose $\bar{\bar{Q}}^*=\aleph_1$. Now as a corollary of (19') we obtain that one of the following statements holds:

(37') (λ) There exists an $x_0\in Q^*$ such that $\overline{\overline{R_2(x,P_l)}}=\aleph_0$ for every $l_0<l<\omega$.

($\lambda\lambda$) There exist sets P, Q and there exists an $l_1<\omega$ such that $P\subseteq P_{l_1}$, $Q\subseteq Q^*$, $\bar{\bar{P}}=\aleph_0$, $\bar{\bar{Q}}=\aleph_1$ and $P\otimes Q\subseteq J_1$.

(Here we apply (19') changing the role of the classes J_1 and J_2.)

Now we prove:

(38') If ($\lambda\lambda$) is satisfied, then either ($\varkappa$) or ($\varkappa\varkappa$) holds.

Let P, Q be sets satisfying ($\lambda\lambda$) for a suitable l_1. We may suppose $P\prec Q$. Since Q is of power $\aleph_1$, $\bar{Q}=\omega_1$ and by the induction hypothesis $\omega_1\to(\omega\cdot n,\omega\cdot 2)$ either ($\varkappa\varkappa$) holds or there exists a subset $S_n\subseteq Q$ such that $[S_n]^2\subseteq J_1$, $\bar{S}_n=\omega\cdot n$.

Put $S_{n+1}=P\cup S_n$. We have $\overline{S_{n+1}}=\omega\cdot(n+1)$ and $[S_{n+1}]^2\subseteq J_1$ by ($\times\times\times\times$), since $[S_{n+1}]^2=[P]^2\cup(P\otimes S_n)\cup[S_n]^2$ and $P\subseteq P_{l_1}$, $P\otimes S_n\subseteq P\otimes Q\subseteq J_1$. Thus we may suppose by (37') and (38') that (λ) holds.

(39') Suppose $P\subseteq P_l$, $\bar{\bar{P}}=\aleph_0$ for an $l<\omega$ and $\bar{\bar{Q}}=\aleph_1$. Using (18') instead of (19') we get here quite similarly to (38') that either one of the statements ($\varkappa$), ($\varkappa\varkappa$) holds, or

[16] Here we use the Ramsey theorem $\aleph_0\to(\aleph_0,\aleph_0)^2$ (see e. g. [1], Theorem 1) and the fact that if for an l_0 and for a subset $S''\subseteq P_{l_0}$, $\bar{\bar{S}}''=\aleph_0$ and $[S_2'']^2\subseteq J_2$, then if S''' is a set such that $\overline{\overline{S'''\cap P_l}}=1$ for an arbitrary $l\neq l_0$, $S'''\subseteq\bigcup_{l\neq l_0} P_l$, then $S'=S''\cup S'''$ satisfies ($\varkappa\varkappa$) and we have nothing to prove.

(μ) there is an element $x_0 \in P$ such that $R_2(x_0, Q) = \aleph_1$,
and so we may suppose that (μ) holds.

Using (λ) and (μ) we are going to define the sequences $\{x_k^1\}_{k<\omega}$, $\{x_k^2\}_{k<\omega}$ by induction on k as follows:[17]

Put $P_l(0) = P_l$ and let Q_0 be a set of power $\aleph_1$ for which $P_l \prec Q_0$ for every $l<\omega$. Now applying (μ) we define x_0^1 in such a way that

$$(40') \qquad x_0^1 \in P_0, \quad \overline{\overline{R_2(x_0^1, Q_0)}} = \aleph_1.$$

Put $\overline{\overline{R_2(x_0, Q_0)}} = U_0$.

By (20′) we may suppose that there exists a subset V_0 such that

$$(41') \qquad V_0 \subseteq U_0 \text{ and } \overline{\overline{R_2(x, V_0)}} = \aleph_1 \text{ for every } x \in V_0.$$

Applying (λ) we can define an x_0^2 satisfying the following conditions:

$$(42') \qquad x_0^2 \in V_0, \quad \overline{\overline{R_2(x_0, P_l)}} = \aleph_0 \text{ for every } 1 \leqq l < \omega.$$

Put $Q_1 = R_2(x_0^2, V_0)$. We have

$$(43') \qquad \overline{\overline{Q}}_1 = \aleph_1, \quad Q_1 \subseteq Q_0.$$

Put further $P_l(1) = R_2(x_0, P_{l+1})$ for every $l<\omega$. It follows from the formulas (40′)—(43′) that the new sequence $P_l(1)$ and $Q(1)$ satisfy the conditions (×)—(××××) and thus we may continue the construction for $k=1$ and so on. So we may suppose that the elements x_k^j as well as the sets Q_k, U_k, V_k, $P_l(k)$ are defined for every $k<\omega, l<\omega, j=1,2$ and that the formulas (40′)—(43′) hold for them with k instead of 0, respectively.

Put $S^1 = \{x_k^1\}_{k<\omega}$, $S^2 = \{x_k^2\}_{k<\omega}$ and $S' = S^1 \cup S^2$. We have

$$(44') \qquad \overline{\overline{S'}} = \omega \cdot 2,$$

since we may suppose $\overline{\overline{S}}^1 = \omega$, $\overline{\overline{S}}^2 = \omega$ and $x_k^1 \in P_k(k) = P_k(0)$ by (40′), $S^2 \subseteq Q_0$ by (42′) and (43′) and we know that $P_k \prec Q_0$ for every $k<\omega$.

This means that the second condition of (κκ) holds. On the other hand, we have by (×××) and (40′)

$$(45') \qquad [S^1]^2 \subseteq J_2.$$

Suppose now that $x_k^1 \in S^1$, $x_{k'}^2 \in S^2$. We have two cases:

(a) $k \leqq k'$, (b) $k > k'$.

If (a) holds, then $x_{k'}^2 \in U_k$ by (41′), (42′) and (43′), hence by (40′) $\{x_k^1, x_k^2\} \in J_2$.

If (b) holds, then $x_k^1 \in P_0(k) \subseteq P_{k'-k+1}(k')$, hence $\{x_k^1, x_{k'}^2\} \in J_2$ in this case too.

[17] The construction used here is similar to that used for the proof of (III).

It follows that

(46′) $$S_1 \otimes S_2 \subseteq J_2 .$$

Let now $x_k^2, x_{k'}^2$ $(k<k')$ be two arbitrary elements of S^2. Then $x_k^2 \in V_k$ by (42′) and $x'_k \in Q_{k'} \subseteq Q_k$ by (43′), hence $\{x_k^2, x_{k'}^2\} \in J_2$.

Therefore we have

(47′) $$[S^2]^2 \subseteq J_2 .$$

Comparing (45′), (46′) and (47′) we get

(48′) $$[S']^2 \subseteq J_2 ,$$

since $[S']^2 = [S^1]^2 \cup S_1 \otimes S_2 \cup [S^2]^2$.

Thus by (44′) and (48′) S' satisfies $(\varkappa\varkappa)$ and Theorem 8 is proved.

4. Almost disjoint sets

(*) Theorem 9. *Suppose $\aleph_\alpha$ is regular. Let S be a set of power $\aleph_{\alpha+1}$. Then there exists a system $\mathfrak{S}$ of subsets of X such that the following conditions hold:*

(i) $\bar{\bar{X}} = \aleph_\alpha$ *for every* $X \in \mathfrak{S}$;

(ii) $\overline{\overline{X \cap Y}} < \aleph_\alpha$ *for every* $X, Y \in \mathfrak{S}$ $(X \neq Y)$;

(iii) *every subset Z of power $\aleph_{\alpha+1}$ of S contains an element of $\mathfrak{S}$.*

Theorem 9 is a consequence of the following more general

(*) Theorem 10. *Suppose $\aleph_\alpha$ is regular. Let S be a set of power $\aleph_{\alpha+1}$, and let $S = \{x_\nu\}_{\nu<\omega_{\alpha+1}}$ be a well-ordering of type $\omega_{\alpha+1}$ of S. Then there exists a system $\mathfrak{S}$ of subsets of S such that the following conditions hold:*

(i′) $\bar{X} = \omega_\alpha$ *for every* $X \in \mathfrak{S}$;

(ii′) $\overline{\overline{X \cap Y}} < \aleph_\alpha$ *for every* $X, Y \in \mathfrak{S}$ $(X \neq Y)$.

If Z is an arbitrary subset of S such that $\bar{Z} = \omega_\alpha^2$, then there exists an $X \in \mathfrak{S}$ for which $X \subseteq Z$.

Here the subsets of S are considered to be ordered by the relation $x_\mu < x_\nu$ if $\mu < \nu$.

Proof. Put $[S]^{\omega_\alpha^2} = \{Z : Z \subseteq S;\ \bar{Z} = \omega_\alpha^2\}$. It is obvious that $[S]^{\omega_\alpha^2} \subseteq [S]^{\aleph_\alpha}$ and thus by (*) we get $\overline{\overline{[S]^{\omega_\alpha^2}}} = \aleph_{\alpha+1}$. Let $[S]^{\omega_\alpha^2} = \{Z_\nu\}_{\nu<\omega_{\alpha+1}}$ be a well-ordering of type $\omega_{\alpha+1}$ of $[S]^{\omega_\alpha^2}$. We are going to define a sequence $\{B_\nu\}_{\nu<\omega_{\alpha+1}}$ of type $\omega_{\alpha+1}$ of the subsets of S by induction on ν as follows:

(1″) Let B_0 be an arbitrary subset of Z_0 such that $\bar{B}_0 = \omega_\alpha$.

Suppose that B_μ is already defined for every $\mu < \nu$ in such a way that $\bar{B}_\mu = \omega_\alpha$.

If $\bar{\nu}<\omega_\alpha$, then the set $Z_\nu-\bigcup_{\mu<\nu} B_\mu$ is of power $\aleph_\alpha$, and we define B_ν so that

$$(2'') \qquad B_\nu \subseteq Z_\nu - \bigcup_{\mu<\nu} B_\mu, \quad \overline{B}_\nu=\omega_\alpha .$$

If $\bar{\nu}=\omega_\alpha$, then let $\{B_\sigma^\nu\}_{\sigma<\omega_\alpha}$ be a well-ordering of type ω_α of the set $\{B_\mu\}_{\mu<\nu}$.

Since the set Z_ν is of type ω_α^2, we can define a sequence $\{Z_{\nu,\varrho}\}_{\varrho<\omega_\alpha}$ so that

$$(3'') \qquad \bigcup_{\varrho<\omega_\alpha} Z_{\nu,\varrho}=Z_\nu, \quad Z_{\nu,\varrho} \prec Z_{\nu,\varrho'} \quad \text{for} \quad \varrho<\varrho', \quad \overline{Z}_{\nu,\varrho}=\omega_\alpha .$$

Since $\overline{B}_\sigma^\nu=\omega_\alpha$ for every $\sigma<\omega_\alpha$, it follows from (3″) that

(4″) there exists at most one $\varrho<\omega_\alpha$ such that $\overline{\overline{B_\sigma^\nu \cap Z_{\nu,\varrho}}}=\aleph_\alpha$.

Now we define a sequence $\{x_\sigma^\nu\}_{\sigma<\omega_\alpha}$ by induction on σ as follows:

(5″) Let x_0^ν be an arbitrary element of $Z_{\nu,0}$.

Suppose that $x_{\sigma'}^\nu$ is defined for every $\sigma'<\sigma$ so that $x_{\sigma'}^\nu \in Z_{\nu,\varrho_{\sigma'}}$ for a $\varrho_{\sigma'}$. Then by (4″) and by the regularity of $\aleph_\alpha$ there exists a ϱ_σ such that $\varrho_\sigma>\varrho_{\sigma'}$ for every $\sigma'<\sigma$ and $\overline{\overline{Z_{\nu,\varrho_\sigma} \cap B_{\sigma'}^\nu}}<\aleph_\alpha$ for every $\sigma'<\sigma$. Using again the regularity of $\aleph_\alpha$ we can define x_σ^ν as an element of $Z_{\nu,\varrho_\sigma}-\bigcup_{\sigma'<\sigma} B_{\sigma'}^\nu$.

Thus the sequence $\{x_\sigma^\nu\}_{\sigma<\omega_\alpha}$ is defined.

Put $B_\nu=\{x_\sigma^\nu\}_{\sigma<\omega_\alpha}$. It follows from (3″) and (5″) that

$$(6'') \qquad \overline{B}_\nu=\omega_\alpha .$$

Thus the sequence B_ν is defined for every $\nu<\omega_{\alpha+1}$ and by (1″), (2″) and (5″) we get

$$(7'') \qquad B_\nu \subseteq Z_\nu \quad \text{for every} \quad \nu<\omega_{\alpha+1} .$$

Suppose now $\mu<\nu$. Then $B_\mu=B_\sigma^\nu$ for a $\sigma<\omega_\alpha$. It follows from (5″) that

$$x_{\sigma'}^\nu \notin B_\sigma^\nu \quad \text{if} \quad \sigma'>\sigma,$$

hence we obtain

$$(8'') \qquad \overline{\overline{B_\mu \cap B_\nu}}<\aleph_\alpha \quad \text{if} \quad \mu \neq \nu .$$

Put $\mathfrak{S}=\{B_\nu\}_{\nu<\omega_{\alpha+1}}$. The system $\mathfrak{S}$ satisfies the requirements of Theorem 10 by (6″), (7″) and (8″). Q. e. d.

PROBLEM 5. Let S be a set of power $\aleph_{\omega+1}$. Does there then exist a system $\mathfrak{S}$ of subsets of S satisfying the requirements (i), (ii), (iii) of Theorem 9 for $\alpha = \omega$?

The author would like to express his thanks to Prof. P. ERDŐS for his helpful criticism.

(Received 1 December 1959)

References

[1] P. ERDŐS and R. RADO, A partition calculus in set theory, *Bull. Amer. Math. Soc.*, **62** (1956), pp. 427—489.

[2] P. ERDŐS and G. FODOR, Some remarks on set theory. V, *Acta Sci. Math. Szeged*, **12** (1956), pp. 250—260.

[3] P. ERDŐS, Some remarks on set theory. III, *Michigan Math. Journal*, **2** (1953), pp. 51—57.

[4] P. ERDŐS and A. HAJNAL, On the structure of set-mappings, *Acta Math. Acad. Sci. Hung.*, **9** (1958), pp. 111—131.

ON AN IMPROVEMENT OF SOME NEW ONE-SIDED THEOREMS OF THE THEORY OF DIOPHANTINE APPROXIMATIONS

By

P. TURÁN (Budapest), member of the Academy

1. In the last years I observed that a reformulation of some theorems of the theory of diophantine approximations in terms of generalized power-sums and subsequent generalization widens considerably the field of applications of the theory in the analysis and analytical number-theory.[1] The applications are based on three main theorems, among which we quote here only the first two, in their latest form.

I. For $n \geqq 2$ and positive integer m, for arbitrary complex b_j's and complex z_j's with

$$|z_1| \geqq |z_2| \geqq \cdots \geqq |z_n| = 1 \tag{1.1}$$

there is an integer ν_0 with

$$m+1 \leqq \nu_0 \leqq m+n$$

such that

$$\left| \sum_{j=1}^{m} b_j z_j^{\nu_0} \right| \geqq \frac{|b_1 + \cdots + b_n|}{\sum_{l=0}^{n-1} 2^l \binom{m+l}{l}} \overset{\text{def}}{=} A\,|b_1 + \cdots + b_n|; \tag{1.2}$$

fixing m, n and an arbitrarily small $\varepsilon > 0$ and replacing the right of (1.2) by $A + \varepsilon$ the inequality is false[2] for a suitable sum

$$\sum_{j=1}^{n} b_j^* z_j^{*\nu}$$

with

$$\min_{j=1,\ldots,n} |z_j^*| = 1.$$

[1] A first systematic treatment of this trend is given in my book *Eine neue Methode in der Analysis und deren Anwendungen*, Akad. Kiadó (Budapest, 1953). A completely rewritten Chinese edition appeared in 1956. An English edition will appear in the Interscience Tracts series.

[2] This sharp form of my first main theorem has been found by E. MAKAI and a little later but independently by N. G. DE BRUIJN. See E. MAKAI, The first main theorem of P. Turán, *Acta Math. Acad. Sci. Hung.*, **10** (1959), pp. 405—411; N. G. DE BRUIJN, On Turán's first main theorem, *ibid.*, **11** (1960), pp. 213—216.

For the applications of this theorem so far my original form

$$\left|\sum_{j=1}^{n} b_j z_j^{\nu_0}\right| \geqq |b_1+\cdots+b_n| \left(\frac{n}{2e(m+n)}\right)^n \tag{1.3}$$

is sufficient.

II. For $n \geqq 2$ and positive integer m, for arbitrary complex b_j's and complex z_j's with

$$1 = |z_1| \geqq |z_2| \geqq \cdots \geqq |z_n| \tag{1.4}$$

there is an integer ν_1 with

$$m+1 \leqq \nu_1 \leqq m+n$$

such that

$$\left|\sum_{j=1}^{n} b_j z_j^{\nu_1}\right| \geqq \left(\frac{n}{8e(m+n)}\right)^n \min_{j=1,\ldots,n} |b_1+\cdots+b_j|; \tag{1.5}$$

replacing the constant $8e$ in (1.5) by any positive $c < \dfrac{2e}{\log 2}$ the inequality becomes false generally.[3]

To illustrate the applicability of the second main theorem I mention the following two recent theorems:

III. Denoting, as usual, the function[4] $\sum_{n \leqq x} \Lambda(n) - x$ by $R(x)$ we have for suitable *explicit* numerical positive constants c_1 and c_2 for $T > c_1$ the inequality

$$\int_{Te^{-\frac{\log T}{\sqrt{\log\log T}}}}^{T} R(x)^2 dx > T^2 e^{-c_2 \frac{\log T \log\log\log T}{\log\log T}} \tag{1.6}$$

from which one easily gets

$$\int_0^T \frac{R(x)^2}{x} dx > T e^{-\frac{1}{2} c_2 \frac{\log T \log\log\log T}{\log\log T}} \tag{1.7}$$

[3] The improved form (1.5) of my original form is contained in our paper with VERA T. Sós, "On some new theorems in the theory of diophantine approximations", *Acta Math. Acad. Sci. Hung.*, **6** (1955), pp. 241–255, apart from a remark of Mr. S. UCHIYAMA; see his paper "A note on the second main theorem of P. Turán", *ibid.*, **9** (1958), pp. 378—380. The ingenious counterexample which shows that $c < \dfrac{2e}{\log 2}$ is false, is due to E. MAKAI; see MAKAI l. c.

[4] $\Lambda(n)$ stands for the Mangoldt-symbol, equals $\log p$ if $n =$ prime power p^α and 0 otherwise.

For the sake of orientation I remark that the inequality

$$\int_0^T \frac{R(x)^2}{x} dx = O(T)$$

has been proved by H. CRAMÉR[5] under supposition of the truth of Riemann's conjecture (while (1. 7) holds without any conjecture).

IV. Denoting for $(l, k) = 1$, as usual, $\sum\limits_{\substack{n \leq x \\ n \equiv l (\mathrm{mod}\, k)}} \Lambda(n)$ by $\psi_l(x, k)$ we have for suitable *explicit* numerical positive constants c_3 and c_4 for $T > \max(c_3, e^{3k^4})$ the inequality[6]

$$\int_{Te^{-\frac{\log T}{\sqrt{\log\log T}}}}^{T} (\psi_1(x, k) - \psi_l(x, k))^2 dx > T^2 e^{-c_4 \frac{\log T \log\log\log T}{\log\log T}}.$$

2. The results concerning the distribution of the sign-changes of $(\pi(x) - \mathrm{li}\, x)$, of $R(x)$ or of $\psi_{l_1}(x) - \psi_{l_2}(x)$ are, in spite of pathbreaking results of E. SCHMIDT, LANDAU, LITTLEWOOD, INGHAM, PÓLYA and SKEWES among others, rather incomplete. Therefore it seems to me desirable to try to refine the two main theorems so that they should be able to cope with such problems, too. This leads at once to the "one-sided" problems which constitute a set of problems of independent interest of the theory; qualitatively speaking they require under conditions (1. 1) or (1. 4) a positive lower bound L_1 for

$$\max_{\substack{m+1 \leq \nu \leq m+10n \\ \nu \text{ integer}}} \mathrm{Re} \sum_{j=1}^{n} b_j z_j^\nu$$

and a negative upper bound L_2 for

$$\min_{\substack{m+1 \leq \nu \leq m+10n \\ \nu \text{ integer}}} \mathrm{Re} \sum_{j=1}^{n} b_j z_j^\nu ,$$

respectively, say, depending only upon m, n and the b_j's. Easy counter-examples show, however, that *generally* (i. e. for independent variable z_j's) such theorems cannot exist. However, the possibility remains to obtain such

[5] See his paper "Some theorems concerning prime numbers", *Arkiv för Math. Astr. och Fys.*, **15** (1921).

[6] The applicability of my methods to the investigation of $\psi_1(x, k) - \psi_l(x, k)$ has been first observed by S. KNAPOWSKI. See his papers "On an explicit lower estimate in prime number theory", *Journal London Math. Soc*, 34 (1959), pp. 437—441; and "On the mean values of certain functions in prime number theory", *Acta Math. Acad. Sci, Hung.*, **10** (1959), pp. 375—390. The results here are slightly stronger.

theorems, imposing certain restrictions upon the variables z_j (apart from (1. 1) or (1.4)). Such restrictions can be twofold; either simple and natural geometrical restrictions which make the problem also interesting in itself, or such ones which are perhaps less simple but which can be verified in the intended applications. As a first example I have chosen the simplest geometrical restriction

$$\varkappa \leqq |\operatorname{arc} z_j| \leqq \pi \qquad (j=1, 2, \ldots, n) \tag{2.1}$$

where

$$0 < \varkappa \leqq \frac{\pi}{2}. \tag{2.2}$$

Under this restriction I have found the following two theorems:[7]

V. For $n \geqq 2$ and positive integer m, for positive b_j's and complex z_j's with (1. 1) and (2. 1) there are integers ν_2 and ν_3 with

$$m+1 \leqq \nu_2 \leqq m+2n\left(1+\frac{\pi}{\varkappa}\right) \tag{2.3}$$

and

$$m+1 \leqq \nu_3 \leqq m+3n\left(1+\frac{\pi}{\varkappa}\right) \tag{2.4}$$

such that

$$\operatorname{Re} \sum_{j=1}^{n} b_j z_j^{\nu_2} \geqq \left(\sum_{j=1}^{n} b_j\right) 16^{-n(m+1)} \tag{2.5}$$

and

$$\operatorname{Re} \sum_{j=1}^{n} b_j z_j^{\nu_3} \leqq -\left(\sum_{j=1}^{n} b_j\right) 4^{-n(2m+1)}. \tag{2.6}$$

In the next theorem only the case $b_1 = b_2 = \cdots = b_n = 1$ is treated which, however, is one of the most important cases in the applications.

VI. For $n \geqq 2$, positive integer m and complex z_j's with (1.4) and (2. 1) there are integers ν_4 and ν_5 with

$$m+1 \leqq \nu_4 \leqq m+2n\left(1+\frac{\pi}{\varkappa}\right) \tag{2.7}$$

and

$$m+1 \leqq \nu_5 \leqq m+3n\left(1+\frac{\pi}{\varkappa}\right) \tag{2.8}$$

[7] See my forthcoming paper "On some one-sided theorems of the theory of diophantine approximations" in the Jubilee volume of *Ind. Math. Soc.*

such that

$$\text{Re} \sum_{j=1}^{n} z_j^{\nu_4} \geqq 8^{-n(n+1)\left\{m+2n\left(1+\frac{\pi}{\varkappa}\right)\right\}} \tag{2.9}$$

and

$$\text{Re} \sum_{j=1}^{n} z_j^{\nu_5} \leqq -8^{-n(n+1)\left\{m+2n\left(1+\frac{\pi}{\varkappa}\right)\right\}}. \tag{2.10}$$

3. These bounds — compared with (1.3) or (1.5) — are very weak, but they constituted the first non-trivial results in this trend. In the present note I shall obtain much better bounds, which are in a certain sense not far from the best-possible, as a comparison with (1.3) and (1.5) shows. We are going to prove the following two theorems:

THEOREM I. *For $n>2$ and positive integer m, for arbitrary complex b_j's with*

$$\text{Re} \sum_{j=1}^{n} b_j > 0 \tag{3.1}$$

for complex z_j's with

$$|z_1| \geqq |z_2| \geqq \cdots \geqq |z_n| = 1 \tag{3.2}$$

nd $\left(0<\varkappa\leqq\frac{\pi}{2}\right)$ with

$$\pi \geqq |\text{arc}\, z_j| \geqq \varkappa \qquad (j=1,\ldots,n) \tag{3.3}$$

there are integers ν_6 and ν_7 with

$$m+1 \leqq \nu_6 \leqq m+n\left(3+\frac{\pi}{\varkappa}\right) \tag{3.4}$$

and

$$m+1 \leqq \nu_7 \leqq m+n\left(3+\frac{\pi}{\varkappa}\right) \tag{3.5}$$

such that the inequalities

$$\text{Re} \sum_{j=1}^{n} b_j z_j^{\nu_6} \geqq \left(\text{Re} \sum_{j=1}^{n} b_j\right) \frac{1}{5n} \left(\frac{n}{27(m+n)}\right)^{2n} \tag{3.6}$$

and

$$\text{Re} \sum_{j=1}^{n} b_j z_j^{\nu_7} \leqq -\left(\text{Re} \sum_{j=1}^{n} b_j\right) \frac{1}{5n} \left(\frac{n}{27(m+n)}\right)^{2n} \tag{3.7}$$

hold.

THEOREM II. *For $n\geqq 2$ and positive integer m, for complex z_j's with*

$$1 = |z_1| \geqq |z_2| \geqq \cdots \geqq |z_n| \tag{3.8}$$

and $\left(0<\varkappa\leqq\frac{\pi}{2}\right)$ with

$$\pi \geqq |\text{arc}\, z_j| \geqq \varkappa \qquad (j=1,\ldots,n) \tag{3.9}$$

there are integers ν_8 and ν_9 with

$$m+1 \leqq \nu_8 \leqq m+n\left(3+\frac{\pi}{\varkappa}\right) \tag{3.10}$$

and

$$m+1 \leqq \nu_9 \leqq m+n\left(3+\frac{\pi}{\varkappa}\right) \tag{3.11}$$

such that the inequalities

$$\operatorname{Re}\sum_{j=1}^{n} z_j^{\nu_8} \geqq \{81(m+n)\}^{-2n^3\left(3+\frac{\pi}{\varkappa}\right)} \tag{3.12}$$

and

$$\operatorname{Re}\sum_{j=1}^{n} z_j^{\nu_9} \leqq -\{81(m+n)\}^{-2n^3\left(3+\frac{\pi}{\varkappa}\right)} \tag{3.13}$$

hold.[8]

Dropping the restriction in Theorem I, m being an integer, we get at once (replacing m by $[m]$)

$$\max_{m \leqq \nu_6 \leqq m+n\left(3+\frac{\pi}{\varkappa}\right)} \operatorname{Re}\sum_{j=1}^{n} b_j z_j^{\nu_6} \geqq \left(\operatorname{Re}\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{n}{27(m+n)}\right)^{2n}. \tag{3.14}$$

Let now in Theorem I b_j be real, let a and d be positive numbers, further let

$$(0<)\,\lambda_1<\lambda_2<\cdots<\lambda_n$$

be such that[9]

$$\left\|\frac{d}{n}\lambda_j\right\| \geqq \varkappa \qquad (j=1,\ldots,n), \tag{3.15}$$

and apply it with

$$z_j = e^{i\lambda_j\frac{d}{n}}, \quad m = a\frac{n}{d}.$$

Then the conditions of the theorem are fulfilled and we get that for suitable integers ν_6 and ν_7 in $a\frac{n}{d} \leqq x \leqq a\frac{n}{d}+n\left(3+\frac{\pi}{\varkappa}\right)$

$$\sum_{j=1}^{n} b_j \cos\lambda_j\left(\frac{d}{n}\nu_6\right) \geqq \left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{1}{27}\frac{d}{a+d}\right)^{2n}$$

[8] Mr. S. Knapowski whom I communicated Theorem II found already an important application of it to the study of the sign of $R(x)$; he intends to publish his results in *Journ. of London Math. Soc.* In the mean time I improved Theorem II considerably; this will be published in the paper "On some further one-sided theorems of new type in the theory of diophantine approximations" in *Acta Math. Acad. Sci. Hung.* The numerous new applications of these results will be treated in a joint paper with Mr. S. Knapowski. *(Added in proof 25 October 1960.)*

[9] $\|x\|$ denotes here the distance of x from the next multiple of 2π.

and

$$\sum_{j=1}^{n} b_j \cos\lambda_j\left(\frac{d}{n}\nu_7\right) \leqq -\left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{1}{27}\frac{d}{a+d}\right)^{2n}.$$

But this gives at once the

COROLLARY. Under condition (3. 15) we have

$$\max_{a\leqq x\leqq a+d\left(3+\frac{\pi}{\varkappa}\right)} \sum_{j=1}^{n} b_j \cos\lambda_j x > \left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{d}{27(a+d)}\right)^{2n}$$

and

$$\min_{a\leqq x\leqq a+d\left(3+\frac{\pi}{\varkappa}\right)} \sum_{j=1}^{n} b_j \cos\lambda_j x < -\left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{d}{27(a+d)}\right)^{2n}.$$

In particular (with $\varkappa=\lambda_1$, $d=n$) : if $(0<)\lambda_1<\cdots<\lambda_n\leqq 2\pi-\lambda_1$, then

$$\max_{a\leqq x\leqq a+n\left(3+\frac{\pi}{\lambda_1}\right)} \sum_{j=1}^{n} b_j \cos\lambda_j x > \left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{n}{27(a+n)}\right)^{2n}$$

and

$$\min_{a\leqq x\leqq a+n\left(3+\frac{\pi}{\lambda_1}\right)} \sum_{j=1}^{n} b_j \cos\lambda_j x < -\left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{n}{27(a+n)}\right)^{2n}.$$

Or still more particularly: if $\frac{\pi}{2}\leqq\lambda_1<\lambda_2<\cdots<\lambda_n\leqq\frac{3\pi}{2}$, then

$$\max_{a\leqq x\leqq a+5n} \sum_{j=1}^{n} b_j \cos\lambda_j x > \left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{n}{27(a+n)}\right)^{2n}$$

and

$$\min_{a\leqq x\leqq a+5n} \sum_{j=1}^{n} b_j \cos\lambda_j x < -\left(\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{n}{27(a+n)}\right)^{2n}.$$

4. First we reproduce shortly the proofs of some lemmas from my Indian paper, for the sake of completeness.

LEMMA I. *If*

(4. 1) $$F(z)=1+a_1z+\cdots+a_Nz^N$$

is a polynomial with real coefficients and with all zeros outside the angle

(4. 2) $$|\text{arc } z|<\varkappa$$

with a $0<\varkappa\leqq\frac{\pi}{2}$, *then there is a polynomial* $\varphi(z)$ *(with real coefficients) such*

that $F(z)\varphi(z)$ *is a polynomial of degree*

$$\leqq \frac{N}{2}\left(1+\left[\frac{\pi}{\varkappa}\right]\right)$$

and with non-negative coefficients.

For the proof we consider first the special case

$$F(z) = F_\alpha(z) \overset{\text{def}}{=} 1-2\cos\alpha\cdot z+z^2$$

with $\varkappa \leqq |\alpha| \leqq \frac{\pi}{2}$. Since for $|z|<1$

$$\frac{1}{1-2\cos\alpha\cdot z+z^2} = \sum_{\nu=0}^{\infty} \frac{\sin(\nu+1)\alpha}{\sin\alpha} z^\nu,$$

we have for an arbitrary positive integer k the identity

$$(1-2z\cos\alpha+z^2)\sum_{\nu=0}^{k} \frac{\sin(\nu+1)\alpha}{\sin\alpha} z^\nu = 1-\frac{\sin(k+2)\alpha}{\sin\alpha} z^{k+1} + \frac{\sin(k+1)\alpha}{\sin\alpha} z^{k+2}.$$

Choosing

$$k = \left[\frac{\pi}{\alpha}\right]-1$$

we have

$$\operatorname{sign}\sin(k+2)\alpha = -1,$$
$$\operatorname{sign}\sin(k+1)\alpha = +1,$$

i. e. for $F_\alpha(z)$ the polynomial

(4. 3) $$\varphi_\alpha(z) \overset{\text{def}}{=} \sum_{\nu=0}^{\left[\frac{\pi}{\alpha}\right]-1} \frac{\sin(\nu+1)\alpha}{\sin\alpha} z^\nu$$

settles the case with

(4. 4) $$F_\alpha(z)\varphi_\alpha(z) = 1+\frac{\left|\sin\left(1+\left[\frac{\pi}{\alpha}\right]\right)\alpha\right|}{\sin\alpha} z^{\left[\frac{\pi}{\alpha}\right]} + \frac{\sin\left[\frac{\pi}{\alpha}\right]\alpha}{\sin\alpha} z^{\left[\frac{\pi}{\alpha}\right]+1}.$$

But (4. 4) allows to settle also the general case. If

$$-\lambda_j \qquad (j=1, 2, \ldots, j_1)$$

are the negative zeros of $F(z)$,

$$r_j e^{\pm i\alpha_j} \qquad (j=1, 2, \ldots, j_2)$$

the complex-conjugate zeros of $F(z)$ with

$$\pi > \alpha_j \geqq \frac{\pi}{2},$$

further

$$\varrho_j e^{\pm i\beta_j} \qquad (j=1,2,\ldots,j_3)$$

the complex-conjugate zeros of $F(z)$ with

$$\frac{\pi}{2} > \beta_j \geqq \varkappa, \tag{4.5}$$

then

$$j_1 + 2j_2 + 2j_3 = N \tag{4.6}$$

and

$$\begin{aligned} F(z) = \prod_{j=1}^{j_1}\left(1+\frac{z}{\lambda_j}\right)\cdot\prod_{j=1}^{j_2}\left\{1-2\cos\alpha_j\cdot\left(\frac{z}{r_j}\right)+\left(\frac{z}{r_j}\right)^2\right\}\cdot \\ \cdot\prod_{j=1}^{j_3}\left\{1-2\cos\beta_j\cdot\left(\frac{z}{\varrho_j}\right)+\left(\frac{z}{\varrho_j}\right)^2\right\} \end{aligned} \tag{4.7}$$

holds. Owing to (4. 4) we may choose as the required $\varphi(z)$ the polynomial[10] (with the notations (4. 3))

$$\sum_{j=1}^{j_3} \varphi_{\beta_j}\left(\frac{z}{\varrho_j}\right). \tag{4.8}$$

The degree of $F(z)\varphi(z)$ is by this choice, using (4. 5) and (4. 6),

$$\begin{aligned} j_1+2j_2+\sum_{j=1}^{j_3}\left(1+\left[\frac{\pi}{\beta_j}\right]\right) \leqq j_1+2j_2+j_3\left(1+\left[\frac{\pi}{\varkappa}\right]\right) \leqq \\ \leqq \left(1+\left[\frac{\pi}{\varkappa}\right]\right)\left(\frac{j_1}{2}+j_2+j_3\right) = \frac{N}{2}\left(1+\left[\frac{\pi}{\varkappa}\right]\right). \end{aligned}$$

Q. e. d.

5. Denoting the coefficients of the above $F(x)\varphi(z)$ by e_ν we need the

LEMMA II. *If the zeros of* $F(z)$ *lie outside the angle* $|\text{arc}\, z| < \varkappa$ *and also outside the circle* $|z|<1$, *then*

$$\sum_\nu e_\nu \leqq 2^N.$$

For the proof of this lemma we remark that

$$\sum_\nu e_\nu = F(1)\varphi(1),$$

[10] Empty product means 1.

i. e. from (4. 7), (4. 8) and (4. 4)

$$\sum_{\nu} e_\nu = \prod_{j=1}^{j_1}\left(1+\frac{1}{\lambda_j}\right)\prod_{j=1}^{j_2}\left(1+2\,|\cos\alpha_j|\,\frac{1}{r_j}+\frac{1}{r_j^2}\right)\cdot$$

$$\cdot\prod_{j=1}^{j_3}\left\{1+\frac{\left|\sin\left(1+\left[\frac{\pi}{\beta_j}\right]\right)\beta_j\right|}{\sin\beta_j}\cdot\frac{1}{\varrho_j}+\frac{\sin\left[\frac{\pi}{\beta_j}\right]\beta_j}{\sin\beta_j}\,\frac{1}{\varrho_j^2}\right\}.$$

Since $\lambda_j \geqq 1$, $r_j \geqq 1$, $\varrho_j \geqq 1$, we have

$$\sum_{\nu} e_\nu \leqq 2^{j_1+2j_2}\prod_{j=1}^{j_3}\left\{1+\frac{\left|\sin\left(1+\left[\frac{\pi}{\beta_j}\right]\right)\beta_j\right|}{\sin\beta_j}+\frac{\sin\left[\frac{\pi}{\beta_j}\right]\beta_j}{\sin\beta_j}\right\}=$$

$$=2^{j_1+2j_2}\prod_{j=1}^{j_3}\left\{1+\frac{-\sin\left(1+\left[\frac{\pi}{\beta_j}\right]\right)\beta_j+\sin\left[\frac{\pi}{\beta_j}\right]\beta_j}{\sin\beta_j}\right\}=$$

$$=2^{j_1+2j_2}\prod_{j=1}^{j_3}\left(1-\frac{\cos\left(\frac{1}{2}+\left[\frac{\pi}{\beta_j}\right]\right)\beta_j}{\cos\frac{\beta_j}{2}}\right)$$

and thus, owing to $(0<)\beta_j<\frac{\pi}{2}$,

$$\sum_{\nu} e_\nu \leqq 2^{j_1+2j_2}(1+\sqrt{2})^{j_3}\leqq 2^N.$$

6. We now prescribe the numbers γ_j with

(6. 1) $$\gamma_j>1 \qquad (j=1,2,\ldots,N-1).$$

Then we shall need the

LEMMA III. *To arbitrary complex* $\xi_1, \xi_2, \ldots, \xi_N$ *with*

$$|\xi_1|\geqq|\xi_2|\geqq\cdots\geqq|\xi_N|>0$$

there is an index k *with* $1\leqq k\leqq N$ *such that the annulus*[11]

$$\frac{|\xi_k|}{\gamma_k}<|z|<|\xi_k|$$

contains none of the ξ_j'*s and in the case* $k>1$ *also the inequality*

$$|\xi_k|\geqq\frac{|\xi_1|}{\gamma_1\gamma_2\ldots\gamma_{k-1}}$$

holds.

[11] In the case $k=N$ the lower bound is 0, of course.

PROOF. If there are indices l with the property

$$|\xi_{l+1}| \leqq \frac{|\xi_l|}{\gamma_l}, \tag{6.2}$$

we choose as k the smallest of them. If $k=1$, we have nothing to prove. If $k>1$, then for all indices v with $1 \leqq v < k$ we have

$$|\xi_{v+1}| > \frac{|\xi_v|}{\gamma_v}$$

and thus

$$|\xi_k| > \frac{|\xi_{k-1}|}{\gamma_{k-1}} > \frac{|\xi_{k-2}|}{\gamma_{k-1}\gamma_{k-2}} > \cdots > \frac{|\xi_1|}{\gamma_{k-1}\gamma_{k-2}\cdots\gamma_1},$$

and, of course, there are no ξ_j's in $\frac{|\xi_k|}{\gamma_k} < |z| < |\xi_k|$. If there are no indices l with the property in (6. 2), this means that

$$|\xi_{v+1}| > \frac{|\xi_v|}{\gamma_v} \qquad (v = 1, 2, \ldots, N-1); \tag{6.3}$$

in this case we choose $k=N$ and (6. 3) gives again

$$|\xi_N| > \frac{|\xi_{N-1}|}{\gamma_{N-1}} > \cdots > \frac{|\xi_1|}{\gamma_{N-1}\gamma_{N-2}\cdots\gamma_1}$$

(and, of course, there are no ξ_v's for $|z| < |\xi_N|$). Q. e. d.

Further we shall use later the following theorem of NÖRLUND:[12]

Let $w_1, w_2, \ldots, w_\nu$ be complex numbers outside a closed rectifiable curve L, let $g(z)$ be analytical outside and on L and vanishing for $z=\infty$. Then writing

$$L_\nu(z, g(z)) = d_0 + d_1(z-w_1) + d_2(z-w_1)(z-w_2) + \cdots$$
$$\cdots + d_{\nu-1}(z-w_1)(z-w_2)\ldots(z-w_{\nu-1})$$

for the (only) polynomial of degree $\leqq \nu-1$ which coincides for $z=w_j$ with $g(w_j)$ $(j=1, 2, \ldots, \nu)$, we have the integral representation

$$d_j = \frac{1}{2\pi i}\int\limits_L \frac{g(w)\,dw}{(w-w_1)(w-w_2)\ldots(w-w_{j+1})} \qquad (j=0, 1, \ldots, \nu-1). \tag{6.4}$$

12 For a verification see e. g. the German edition of my book, p. 43.

7. Next we turn to the proof of our Theorem I. We introduce the numbers ϑ_j $(j=1,2,\dots,2n)$ by

$$\left.\begin{aligned} \vartheta_j &= z_j \qquad (7.1)\\ \vartheta_{n+j} &= \bar{z}_j \qquad (7.2)\end{aligned}\right\} \quad \text{for} \quad j=1,2,\dots,n;$$

hence we have

$$(7.3)\qquad \begin{gathered} |\vartheta_1| \geqq \vartheta_2| \geqq \cdots \geqq |\vartheta_n| = 1, \\ \varkappa \leqq |\operatorname{arc} \vartheta_j| \leqq \pi \quad \text{for} \quad j=1,2,\dots,2n. \end{gathered}$$

Among the ϑ_j's there are possibly equal ones; let $\eta_1, \eta_2, \dots, \eta_l$ be the maximal numbers of different ones among the ϑ_j's with $|\eta_1| \geqq |\eta_2| \geqq \cdots \cdots \geqq |\eta_l| = 1$, say; evidently $l \geqq 2$. Let

$$(7.4)\qquad \begin{aligned} L_l(z, z^{-m-1}) = d_0 + d_1(z-\eta_1) + d_2(z-\eta_1)(z-\eta_2) + \cdots \\ \cdots + d_{l-1}(z-\eta_1)(z-\eta_2)\dots(z-\eta_{l-1}) \end{aligned}$$

be the polynomial assuming for $z=\eta_\nu$ the value η_ν^{-m-1} $(\nu=1,2,\dots,l)$. Since $|\eta_\nu| \geqq 1$, we may choose as contour L the circle

$$(7.5)\qquad |z| = 1 - \frac{1}{2}\frac{n}{m+n}$$

and apply NÖRLUND's formula in (6.4) with $g(z) = z^{-m-1}$. This gives

$$(7.6)\qquad d_j = \frac{1}{2\pi i} \int\limits_{|z|=1-\frac{1}{2}\frac{n}{m+n}} \frac{dz}{z^{m+1}(z-\eta_1)(z-\eta_2)\dots(z-\eta_{j+1})} \quad (j=0,1,\dots,l-1).$$

Since we shall need the polynomial $L_l(z, z^{-m-1})$ in the form

$$(7.7)\qquad c_0^{(1)} + c_1^{(1)} z + \cdots + c_{l-1}^{(1)} z^{l-1},$$

(7.4) and (7.6) give

$$(7.8)\qquad c_{l-1}^{(1)} = d_{l-1} = \frac{1}{2\pi i} \int\limits_{|z|=1-\frac{1}{2}\frac{n}{m+n}} \frac{dz}{z^{m+1}(z-\eta_1)\dots(z-\eta_l)}$$

and for $0 \leqq j \leqq l-2$.

$$c_j^{(1)} = d_j - d_{j+1} \sum_{1 \leqq j_1 \leqq j+1} \eta_{j_1} + d_{j+2} \sum_{1 \leqq j_1 < j_2 \leqq j+2} \eta_{j_1}\eta_{j_2} - \cdots =$$

$$= \frac{1}{2\pi i} \int\limits_{|z|=1-\frac{1}{2}\frac{n}{m+n}} \frac{1}{z^{m+1}(z-\eta_1)\dots(z-\eta_{j+1})} \left\{ 1 - \frac{\sum\limits_{1 \leqq j_1 \leqq j+1} \eta_{j_1}}{z-\eta_{j+2}} + \right.$$

$$\left. + \frac{\sum\limits_{1 \leqq j_1 < j_2 \leqq j+2} \eta_{j_1}\eta_{j_2}}{(z-\eta_{j+2})(z-\eta_{j+3})} - \cdots \right\} dz$$

or

$$|c_j^{(1)}| \leqq \frac{1}{2\pi} \int\limits_{|z|=1-\frac{1}{2}\frac{n}{m+n}} \frac{|dz|}{|z|^{m+1}} \left\{ \frac{1}{|z-\eta_1|\,|z-\eta_2|\ldots|z-\eta_{j+1}|} + \right. \tag{7.9}$$

$$\left. + \frac{\sum\limits_{1\leqq j_1\leqq j+1} |\eta_{j_1}|}{|z-\eta_1|\,|z-\eta_2|\ldots|z-\eta_{j+2}|} + \frac{\sum\limits_{1\leqq j_1<j_2\leqq j+2} |\eta_{j_1}|\,|\eta_{j_2}|}{|z-\eta_1|\,|z-\eta_2|\ldots|z-\eta_{j+3}|} + \cdots \right\}.$$

We consider the terms

$$\frac{1}{|z-\eta_1|\,|z-\eta_2|\ldots|z-\eta_{j+1}|} \overset{\text{def}}{=} U_{j,0}$$

and

$$\frac{\sum\limits_{1\leqq j_1<j_2<\cdots<j_s\leqq j+s} |\eta_{j_1}|\,|\eta_{j_2}|\ldots|\eta_{j_s}|}{|z-\eta_1|\,|z-\eta_2|\ldots|z-\eta_{j+s+1}|} \overset{\text{def}}{=} U_{js} \tag{7.10}$$

in (7.9) for $s=1,2,\ldots,l-j-1$. If all η_j's are absolutely $\leqq 2$, then

$$U_{js} \leqq \frac{\binom{j+s}{s} 2^s}{\left(\frac{1}{2}\frac{n}{m+n}\right)^{j+s+1}} \qquad (s=0,1,\ldots,l-j-1)$$

and thus from (7.9)

$$|c_j^{(1)}| \leqq \left(\frac{1}{1-\frac{1}{2}\frac{n}{m+n}}\right)^{m+1} \sum_{s=0}^{l-j-1} \binom{j+s}{s} 2^s \left(2\frac{m+n}{n}\right)^{j+s+1} < e^n \left(4\frac{m+n}{n}\right)^l \binom{l}{l-j-1}. \tag{7.11}$$

If not all η's are absolutely $\leqq 2$, then there is an index λ such that

$$|\eta_1| \geqq |\eta_2| \geqq \cdots \geqq |\eta_{l-\lambda}| > 2 \geqq |\eta_{l-\lambda+1}| \geqq \cdots \geqq |\eta_{l-1}| = |\eta_l| = 1.$$

Let us observe that for $h \leqq l-\lambda$ and $|z| = 1-\frac{1}{2}\frac{n}{m+n}$ we have

$$\frac{|\eta_h|}{|z-\eta_h|} < \frac{|\eta_h|}{|\eta_h|-1} \leqq 2 < 4\frac{m+n}{n}$$

and for $h > l-\lambda$ and $|z| = 1-\frac{1}{2}\frac{n}{m+n}$

$$\frac{|\eta_h|}{|z-\eta_h|} \leqq \frac{2}{\frac{1}{2}\frac{n}{m+n}} = 4\frac{m+n}{n}.$$

Hence owing to $|\eta_h| \geqq 1$ for $s=0,1,\ldots,l-j+1$

$$U_{js} \leqq \binom{j+s}{s} \prod_{h=1}^{j+s-1} \frac{|\eta_h|}{|z-\eta_h|}$$

and thus

$$U_{js} \leqq \binom{j+s}{s}\left(4\frac{m+n}{n}\right)^{j+s+1}$$

But then from this and (7.9)

$$|c_j^{(1)}| \leqq \left(\frac{1}{1-\frac{1}{2}\frac{n}{m+n}}\right)^{m+1} \sum_{s=0}^{l-j-1} U_{js} \leqq e^n \sum_{s=0}^{l-j-1} \binom{j+s}{s}\left(4\frac{m+n}{n}\right)^{j+s+1} <$$

$$< e^n\left(4\frac{m+n}{n}\right)^l \binom{l}{l-j-1}.$$

This and (7.11) give for all $0 \leqq j \leqq l-1$

$$|c_j^{(1)}| \leqq e^n\left(4\frac{m+n}{n}\right)^l \binom{l}{j+1}. \tag{7.12}$$

8. With the above η's we form the polynomial

$$\omega(z) = \prod_{j=1}^{l}\left(1-\frac{z}{\eta_j}\right). \tag{8.1}$$

Applying Lemma I to $F(z)=\omega(z)$ and $N=l$ we have a polynomial $\varphi(z)$ such that the degree of $\omega(z)\varphi(z)$ is

$$\leqq \frac{l}{2}\left(1+\left[\frac{\pi}{\varkappa}\right]\right),$$

its coefficients are non-negative with the sum $\leqq 2^l$, owing to Lemma II. We now form the polynomial

$$G(z) = 2\left(8\frac{m+n}{n}\right)^l e^n \omega(z)\varphi(z)\,(1+z+z^2+\cdots+z^{l-1}). \tag{8.2}$$

The degree of $G(z)$ is at most $\frac{l}{2}\left(3+\left[\frac{\pi}{\varkappa}\right]\right)-1$, its coefficients are non-negative; moreover, since

$$\omega(z)\varphi(z) = 1+\cdots,$$

the coefficients of $z^0, z, z^2, \ldots, z^{l-1}$ in $\omega(z)\varphi(z)(1+z+\cdots+z^{l-1})$ are at least 1 and hence from (8.2)

$$\text{coeffs. } z^j \text{ in } G(z) \geqq 2\left(8\frac{m+n}{n}\right)^l e^n \qquad (j=0,1,\ldots,l-1). \tag{8.3}$$

With this $G(z)$ and $L_l(z, z^{-m-1})$ we form finally the polynomial

(8.4) $$Q(z)=L_l(z, z^{-m-1})+G(z)=\sum c_\nu^{(2)} z^\nu$$

of degree $\leqq \frac{l}{2}\left(3+\left[\frac{\pi}{\varkappa}\right]\right)-1$. Let us observe from (7.7), (7.12), (8.3) and (8.4) that

(8.5) $$c_\nu^{(2)} \geqq 0 \qquad (\nu=0, 1, \ldots, l-1)$$

and the same inequality holds evidently for

$$l \leqq \nu \leqq \frac{l}{2}\left(3+\left[\frac{\pi}{\varkappa}\right]\right)-1.$$

Owing to the structure of $Q(z)$ we have

$$\sum_\nu c_\nu^{(2)} \eta_j^\nu = L_l(z, z^{-m-1})_{z=\eta_j} = \frac{1}{\eta_j^{m+1}},$$

i. e. for $j=1, 2, \ldots, 2n$

(8.6) $$\sum_{0 \leqq \nu \leqq \frac{l}{2}\left(3+\left[\frac{\pi}{\varkappa}\right]\right)-1} c_\nu^{(2)} \vartheta_j^{m+\nu+1} = 1.$$

If

(8.7) $$b_j = \bar{b}_{j-n} \quad \text{for} \quad j=n+1, \ldots, 2n,$$

multiplying in (8.6) by b_j, summing for $j=1, 2, \ldots, 2n$ we get

(8.8) $$\sum_\nu c_\nu^{(2)} \left(\operatorname{Re} \sum_{j=1}^n b_j z_j^{m+\nu+1}\right) = \operatorname{Re} \sum_{j=1}^n b_j.$$

Owing to the non-negativity of the $c_\nu^{(2)}$'s it follows, taking also $l \leqq 2n$ into account,

(8.9) $$\max_{m+1 \leqq \nu \leqq m+n\left(3+\frac{\pi}{\varkappa}\right)} \operatorname{Re} \sum_{j=1}^n b_j z_j^\nu \geqq \frac{\operatorname{Re} \sum_{j=1}^n b_j}{\sum_\nu c_\nu^{(2)}}.$$

We still need an upper bound for $c_\nu^{(2)}$. (8.4) and (8.2) give

(8.10) $$\sum_\nu c_\nu^{(2)} = L_l(z, z^{-m-1})_{z=1} + G(1) = \sum_\nu c_\nu^{(1)} + 2le^n\left(8\frac{m+n}{n}\right)^l \omega(1)\varphi(1) \leqq$$
$$\leqq \sum_\nu |c_\nu^{(1)}| + 2le^n\left(16\frac{m+n}{n}\right)^l$$

owing to Lemma II (with $F=\omega$). Since from (7.12) and (8.10) we have

$$\sum c_\nu^{(2)} \leqq \left(8\frac{m+n}{n}\right)^l e^n + 2le^n\left(16\frac{m+n}{n}\right)^l \leqq$$

$$\leqq (4n+1)\left(16\sqrt{e}\,\frac{m+n}{n}\right)^{2n} < 5n\left(27\frac{m+n}{n}\right)^{2n},$$

(8.9) gives

$$(8.11)\qquad \max_{\substack{m+1\leqq\nu\leqq m+n\left(3+\frac{\pi}{\varkappa}\right)\\ \nu\ \text{integer}}} \operatorname{Re}\sum_{j=1}^{n} b_j z_j^\nu \geqq \left(\operatorname{Re}\sum_{1}^{n} b_j\right)\frac{1}{5n}\left(\frac{n}{27(m+n)}\right)^{2n},$$

which proves the first part of Theorem I.

9. Next we turn to the second part of Theorem I; the necessary changes are slight. We consider instead of $Q(z)$ in (8.4)

$$(9.1)\qquad Q_1(z) = -L(z, z^{-m-1}) + G(z) = \sum_{\nu\leqq\frac{l}{2}\left(3+\left[\frac{\pi}{\varkappa}\right]\right)-1} c_\nu^{(3)} z^\nu .$$

Owing to (7.12) and (8.3), as above, we get

$$(9.2)\qquad c_\nu^{(3)} \geqq 0 \qquad \left(\nu \leqq \frac{l}{2}\left(3+\left[\frac{\pi}{\varkappa}\right]\right)-1\right).$$

With the above η's, taking into account the form $G(z)$ in (8.2) and also (9.1), we get

$$\sum_\nu c_\nu^{(3)}\eta_j^\nu = Q_1(\eta_j) = -L_l(z, z^{-m-1})_{z=\eta_j} = -\frac{1}{\eta_j^{m+1}}$$

or

$$(9.3)\qquad \sum_\nu c_\nu^{(3)}\vartheta_j^{m+\nu+1} = -1 \qquad (j=1,2,\ldots,2n).$$

With the convention (8.7) this gives

$$\sum_\nu c_\nu^{(3)}\left(\operatorname{Re}\sum_{j=1}^{n} b_j z_j^{m+\nu+1}\right) = -\operatorname{Re}\sum_{j=1}^{n} b_j$$

and owing to the non-negativity of the $c_\nu^{(3)}$'s

$$(9.4)\qquad \min_{\substack{m+1\leqq\nu\leqq m+n\left(3+\frac{\pi}{\varkappa}\right)\\ \nu\ \text{integer}}} \operatorname{Re}\sum_{j=1}^{n} b_j z_j^\nu \leqq -\frac{\operatorname{Re}\sum_{j=1}^{n} b_j}{\sum_\nu c_\nu^{(3)}}.$$

From (9. 1) we get as in (8. 10)

$$\sum_{\nu} c_\nu^{(8)} = -\sum_{\nu=0}^{l-1} c_\nu^{(1)} + 2le^n\left(8\frac{m+n}{n}\right)^l \omega(1)\varphi(1) \leqq \sum_{\nu=0}^{l-1}|c_\nu^{(1)}| + 2le^n\left(16\frac{m+n}{n}\right)^l <$$

(9. 5)

$$< 5n\left(27\frac{m+n}{n}\right)^{2n}$$

and thus

$$\min_{\substack{m+1\leqq\nu\leqq m+n\left(3+\frac{\pi}{\varkappa}\right)\\ \nu\,\text{integer}}} \operatorname{Re}\sum_{j=1}^{n} b_j z_j^\nu \leqq -\left(\operatorname{Re}\sum_{j=1}^{n} b_j\right)\frac{1}{5n}\left(\frac{n}{27(m+n)}\right)^{2n},$$

which proves the second part of Theorem I.

10. Next we turn to the proof of Theorem II. We apply Lemma III with $N=n$, $\xi_\nu = z_\nu$ and

$$\gamma_\nu = \left\{10n^2\left(27\frac{m+\nu}{\nu}\right)^{2\nu}\right\}^{\frac{1}{m+1}}. \tag{10. 1}$$

This assures the existence of a k $(1\leqq k\leqq n)$ such that there is no z_ν in the annulus

$$\frac{|z_k|}{\gamma_k} < |z| < |z_k| \tag{10. 2}$$

and in the case of $k>1$ the estimation

$$|z_k| \geqq \frac{1}{\gamma_1\gamma_2\cdots\gamma_{k-1}} \tag{10. 3}$$

holds. Applying Theorem I to $\frac{z_1}{|z_k|}, \frac{z_2}{|z_k|}, \ldots, \frac{z_k}{|z_k|}$ and

$$b_1 = b_2 = \cdots = b_k = 1$$

we get the existence of integers ν_8 and ν_9 with

$$m+1 \leqq \nu_8 \leqq m+k\left(3+\frac{\pi}{\varkappa}\right)\left(\leqq m+n\left(3+\frac{\pi}{\varkappa}\right)\right)$$

and

$$m+1 \leqq \nu_9 \leqq m+k\left(3+\frac{\pi}{\varkappa}\right)\left(\leqq m+n\left(3+\frac{\pi}{\varkappa}\right)\right)$$

such that

$$\operatorname{Re}\sum_{j=1}^{k} z_j^{\nu_8} \geqq \frac{1}{5k}\left(\frac{k}{27(m+k)}\right)^{2k}|z_k|^{\nu_8} \tag{10. 4}$$

and

$$\operatorname{Re}\sum_{j=1}^{k} z_j^{\nu_9} \leqq -\frac{1}{5k}\left(\frac{k}{27(m+k)}\right)^{2k}|z_k|^{\nu_9}. \tag{10. 5}$$

If $k<n$, then owing to (3.9) and (10.2) we have

$$\left|\mathrm{Re}\sum_{j=k+1}^{n} z_j^{\nu}\right| < n|z_{k+1}|^{\nu} \leqq n\frac{|z_k|^{\nu}}{\gamma_k^{\nu}}$$

and thus

$$\text{(10.6)}\quad \begin{aligned}\mathrm{Re}\sum_{j=1}^{n} z_j^{\nu_8} &\geqq \left\{\frac{1}{5k}\left(\frac{k}{27(m+k)}\right)^{2k} - \frac{n}{\gamma_k^{\nu_8}}\right\}|z_k|^{\nu_8} \geqq \\ &\geqq \left\{\frac{1}{5k}\left(\frac{k}{27(m+k)}\right)^{2k} - \frac{n}{\gamma_k^{m+1}}\right\}|z_k|^{\nu_8} > \frac{1}{10k}\left(\frac{k}{27(m+k)}\right)^{2k}|z_k|^{\nu_8},\end{aligned}$$

using also (10.1). Similarly from (10.5)

$$\text{(10.7)}\quad \mathrm{Re}\sum_{j=1}^{n} z_j^{\nu_9} \leqq -\frac{1}{10k}\left(\frac{k}{27(m+k)}\right)^{2k}|z_k|^{\nu_9}.$$

If $k=1$, this gives

$$\text{(10.8)}\quad \mathrm{Re}\sum_{j=1}^{n} z_j^{\nu_8} \geqq \frac{1}{10}\left(\frac{1}{27(m+1)}\right)^{2}$$

and

$$\text{(10.9)}\quad \mathrm{Re}\sum_{j=1}^{n} z_j^{\nu_9} \leqq -\frac{1}{10}\left(\frac{1}{27(m+1)}\right)^{2}.$$

If, further, $1<k<n$, it follows from (10.1) and (10.3)

$$\begin{aligned}|z_k|^{\nu_8} &\geqq \left(\frac{1}{\gamma_1\gamma_2\cdots\gamma_{k-1}}\right)^{m+n\left(3+\frac{\pi}{\varkappa}\right)} = \prod_{j=1}^{k-1}\left\{\frac{1}{10n^2}\left(\frac{j}{27(m+j)}\right)^{2j}\right\}^{\frac{m+n\left(3+\frac{\pi}{\varkappa}\right)}{m+1}} > \\ &> \prod_{j=1}^{n}\left\{\frac{1}{10n^2}\left(\frac{j}{27(m+j)}\right)^{2j}\right\}^{\frac{m+n\left(3+\frac{\pi}{\varkappa}\right)}{m+1}} > \left\{(4n)^{-2n}\left(\frac{1}{27(m+n)}\right)^{n(n+1)}\right\}^{n\left(3+\frac{\pi}{\varkappa}\right)} > \\ &> \left(\frac{1}{81(m+n)}\right)^{n^2(n+1)\left(3+\frac{\pi}{\varkappa}\right)}\end{aligned}$$

and thus from (10.6)

$$\mathrm{Re}\sum_{j=1}^{n} z_j^{\nu_8} > \frac{1}{10n}\left(\frac{1}{27(m+n)}\right)^{2n}\left(\frac{1}{81(m+n)}\right)^{n^2(n+1)\left(3+\frac{\pi}{\varkappa}\right)} \geqq \left(\frac{1}{81(m+n)}\right)^{2n^3\left(3+\frac{\pi}{\varkappa}\right)}.$$

The same reasoning (even still simpler) holds for $k=n$. Since $|z_k|^{\nu_9}$ can be estimated from below exactly in the same way, the proof of Theorem II is finished.

(Received 3 December 1959)

APPLICATION D'UNE NOUVELLE MÉTHODE DE SOMMATION AUX SÉRIES TRIGONOMÉTRIQUES ET DE DIRICHLET

Par

M. MIKOLÁS (Budapest)

(*Présenté par* P. TURÁN)

En commémoration du professeur LÉOPOLD FEJÉR, *décédé le 15 octobre 1959*

§ 1. Introduction

1. On sait que plusieurs méthodes de sommation se sont montrées avantageuses dans la théorie des séries trigonométriques depuis 1900, année de la découverte du théorème fondamental de FEJÉR.[1] Il est connu qu'au cas des procédés classiques (tout d'abord de ceux de Cesàro et d'Abel—Poisson), la sommabilité en un point x de la série de Fourier d'une fonction $f(u) \in L(0, 2\pi)$ dépend de la nature de cette fonction au voisinage infinitésimal *bilatéral* de x; en particulier, la somme $\frac{1}{2}[f(x+0)+f(x-0)]$ s'obtient pourvu que x soit un point "régulier".[2]

2. Dans cet article, nous allons traiter une nouvelle méthode de sommation des séries trigonométriques, qui est liée avec une théorie récente des intégrales et dérivées d'ordre complexe [9] (s'appuyant sur une généralisation de l'intégrale fractionnaire de WEYL). Cette "méthode $(\mathfrak{W}_{\pm})$", appliquée d'abord à la *série de Fourier* d'une fonction $f(u)$ de période 1, bornée et mesurable, nous amène à une condition simple — *nécessaire et suffisante* — pour la sommabilité en un point x, notamment à l'existence de

$$f\langle x \pm 0\rangle = \lim_{\theta \to +0} \left(\theta \int_0^{\delta} f(x \pm t) t^{\theta-1}\, dt \right), \tag{1.1}$$

$\delta > 0$ étant fixé et arbitrairement petit.[3] Pour la méthode $(\mathfrak{W}_+)$ ou $(\mathfrak{W}_-)$, il s'agit donc d'une condition *unilatérale* en ce sens que les valeurs de $f(u)$

[1] Remarquons que, contre la plupart des citations, ce théorème a été publié dans [1] (et non pas dans [2]) pour la première fois. Voir aussi [7], p. 36.

[2] Cela s'explique évidemment par le fait, que chacune des deux extrémités de $[0, 2\pi]$ est un point "exceptionnel" pour les intégrales singulières correspondantes.

[3] Il ressort que, si la limite (1. 1) existe pour une valeur particulière de δ, elle est indépendante de δ; $f\langle x \pm 0\rangle$ donne alors, respectivement, la somme $(\mathfrak{W}_+)$ ou $(\mathfrak{W}_-)$ de la série de Fourier.

respectivement dans $(x, x+\delta)$ et $(x-\delta, x)$ y figurent seulement.[4] En particulier, la somme $(\mathfrak{W}_+)$ est $f(x+0)$, la somme $(\mathfrak{W}\ )$ est $f(x-0)$, pourvu que la limite en question existe; dans tout segment de continuité, la sommabilité $(\mathfrak{W}_\pm)$ est uniforme. (Cf. Théorème I.) "L'effectivité" de la méthode peut être illustrée en discutant la relation de sommabilité $(\mathfrak{W}_\pm)$ et des *points de Lebesgue:* on trouve par exemple que l'existence de la limite de $h^{-1}\int_0^h f(x+t)dt$ pour $h\to+0$ entraîne celle de $f\langle x+0\rangle$ (analoguement pour les limites correspondantes à gauche), mais $f\langle x+0\rangle, f\langle x-0\rangle$ peuvent exister simultanément sans que

$$(1.2) \qquad \lim_{h\to 0}\left(\frac{1}{2h}\int_{x-h}^{x+h} f(t)dt\right)$$

existe (Théorème II).

Par ailleurs, le Théorème I fournit en même temps une solution alternative d'un *problème dû à* L. Fejér (cf. [3], [10]): déterminer séparément les limites $f(x+0), f(x-0)$ de la série de Fourier de $f(u)$ et par là la différence (saut) $f(x+0)-f(x-0)$, x désignant un point régulier.

3. L'extension (1. 1) des limites unilatérales peut aussi s'utiliser lorsqu'on applique la sommation $(\mathfrak{W}_\pm)$ aux séries de Fourier dérivées (cf. § 5). En ce lieu, nous avons besoin de certaines généralisations $f_\pm^{[p]}(x)$ des dérivées d'ordre p à droite et à gauche $(p=0, 1, \ldots)$; en particulier, $f_\pm^{[0]}(x)=f\langle x\pm 0\rangle$ et $f_-^{[p]}(x)$ est une "limite W_{-p}" au sens de [9].

Enfin — ce qui est une extension naturelle — nous considérons les séries de la forme (s complexe)

$$(1.3) \qquad \begin{cases} \sum_{n=1}^{\infty} 2(2n\pi)^{-s}\left[\alpha_n \cos\left(2n\pi x-\frac{\pi s}{2}\right)+\beta_n \sin\left(2n\pi x-\frac{\pi s}{2}\right)\right], \\ \alpha_n=\int_0^1 f(t)\cos 2n\pi t\, dt, \quad \beta_n=\int_0^1 f(t)\sin 2n\pi t\, dt, \end{cases}$$

à condition que $f(u)$ ou l'une quelconque de ses dérivées appartienne à une classe $L^q(0, 1)$ $(1<q<\infty)$; le résultat nous permettra (par exemple) de sommer des *séries de Dirichlet* ordinaires sur la frontière de leur domaine de convergence dans des cas "difficiles", la somme étant exprimée par le moyen des *limites* $W_s\ f_{[s]}(x)$. (Cf. Théorème III.)

[4] La raison de ce fait est, comme nous verrons, que le "noyau" $\mathfrak{Z}_\theta(u)=\Gamma(\theta)^{-1}\cdot\zeta(1-\theta, u)$ $(0<\theta<1, 0<u<1)$ a un seul point singulier dans $[0, 1]$, à savoir $u=0$.

§ 2. La méthode $(\mathfrak{W}_{\pm})$ de sommation; la fonction auxiliaire $\mathfrak{Z}_s(u)$

1. Partons d'une série trigonométrique

$$A_0+2\sum_{n=1}^{\infty}(A_n\cos 2n\pi u+B_n\sin 2n\pi u), \tag{2.1}$$

A_n, B_n désignant des nombres réels ou complexes arbitraires et u une variable réelle.

Les intégrales (les plus simples) d'ordre p de $\cos 2n\pi x$, $\sin 2n\pi x$ étant de la forme $(2n\pi)^{-p}\cos\left(2n\pi u-\frac{p\pi}{2}\right)$ et $(2n\pi)^{-p}\sin\left(2n\pi u-\frac{p\pi}{2}\right)$ ($p>0$, entier), il s'impose de considérer au lieu de la limite des sommes partielles de (2. 1) les limites:

$$A_0+\lim_{\theta\to+0}\sum_{n=1}^{\infty}\frac{2}{(2n\pi)^{\theta}}\left[A_n\cos\left(2n\pi u\pm\frac{\pi\theta}{2}\right)+B_n\sin\left(2n\pi u\pm\frac{\pi\theta}{2}\right)\right]. \tag{2.2}$$

DÉFINITION. *Supposons la convergence de*

$$\sum_{n=1}^{\infty}2(2n\pi)^{-\theta}\left[A_n\cos\left(2n\pi x+\frac{\pi\theta}{2}\right)+B_n\sin\left(2n\pi x+\frac{\pi\theta}{2}\right)\right]=\sum\nolimits_+(x,\theta) \tag{2.3}$$

pour $\theta>0$ suffisamment petit.

Si la limite de $\sum_+(x,\theta)$ existe pour $\theta\to+0$, on dit que la série (2.1) *est sommable $(\mathfrak{W}_+)$ au point x avec la somme $(\mathfrak{W}_+)$*

$$A_0+\lim_{\theta\to+0}\sum\nolimits_+(x,\theta). \tag{2.4}$$

La sommabilité $(\mathfrak{W}_-)$ et la somme $(\mathfrak{W}_-)$ sont définies d'une façon analogue, en remplaçant (2. 3) *par*

$$\sum_{n=1}^{\infty}2(2n\pi)^{-\theta}\left[A_n\cos\left(2n\pi x-\frac{\pi\theta}{2}\right)+B_n\sin\left(2n\pi x-\frac{\pi\theta}{2}\right)\right]=\sum\nolimits_-(x,\theta). \tag{2.5}$$

Remarquons que, en considérant (2. 1) sous la forme

$$\sum_{n=-\infty}^{\infty}C_n e^{2n\pi i u}\qquad (C_0=A_0;\ C_n=A_{|n|}-i(\operatorname{sg} n)B_{|n|},\ n\neq 0),$$

on obtient pour $\sum_{\pm}(x,\theta)$ les expressions

$$\sum_{n=-\infty}^{\infty}{}' C_n(\mp 2n\pi i)^{-\theta}e^{2n\pi i u},$$

avec la valeur principale de la puissance d'exposant $-\theta$ et avec la restriction $n\neq 0$.

2. Tout d'abord, nous énonçons le simple

LEMME. *Si* (2. 1) *et sa série conjuguée convergent simultanément en un point* x,[5] *alors* (2. 1) *est sommable* ($\mathfrak{W}_\pm$) *en ce point, les sommes* ($\mathfrak{W}_+$) *et* ($\mathfrak{W}_-$) *étant identiques à la somme ordinaire de* (2.1).

En effet, dans notre condition on peut écrire

$$(2.6)\qquad \begin{cases} \sum_\pm(x,\theta) = \cos\frac{\pi\theta}{2}\sum_{n=1}^{\infty}\frac{2}{(2n\pi)^\theta}(A_n\cos 2n\pi x + B_n\sin 2n\pi x)\pm \\ \qquad \pm\sin\frac{\pi\theta}{2}\sum_{n=1}^{\infty}\frac{2}{(2n\pi)^\theta}(B_n\cos 2n\pi x - A_n\sin 2n\pi x), \end{cases}$$

ce qui fournit immédiatement pour $\theta\to+0$ le résultat en question par une proposition "de type Abel" sur les séries de Dirichlet (cf. [4]).

Il est clair que la condition de la convergence de la série conjuguée peut être remplacée par la restriction plus faible

$$\sum_{n=1}^{\infty} 2(2n\pi)^{-\theta}(B_n\cos 2n\pi x - A_n\sin 2n\pi x) = O(1) \qquad (\theta\to+0).$$

3. Pour obtenir un exemple non trivial, posons $A_0=0$, $A_n=B_n=1$ ($n>1$). Alors on aura, par une formule importante de HURWITZ:[6]

$$(2.7)\qquad \begin{cases} \sum_\pm(x,\theta) = \sum_{n=1}^{\infty}\frac{2}{(2n\pi)^\theta}\cos\left(2n\pi x\mp\frac{\pi\theta}{2}\right) = \\ = \cos\frac{\pi\theta}{2}\sum_{n=1}^{\infty}\frac{2\cos 2n\pi x}{(2n\pi)^\theta}\mp\sin\frac{\pi\theta}{2}\sum_{n=1}^{\infty}\frac{2\sin 2n\pi x}{(2n\pi)^\theta} = \\ = \Gamma(\theta)^{-1}\bar\zeta(1-\theta,\mp x) \qquad (0<\theta<1,\ 0<x<1), \end{cases}$$

où $\bar\zeta(s,u)$ signifie une fonction telle que $\bar\zeta(s,u+1)=\zeta(s,u)$ et $\bar\zeta(s,u)=\zeta(s,u)$ $(0<u\leqq 1)$, $\zeta(s,u)$ désignant la fonction ζ généralisée.[7]

Pour plus de brièveté, soit

$$(2.8)\qquad \mathfrak{Z}_s(u) = \begin{cases} \Gamma(s)^{-1}\bar\zeta(1-s,u) & (s\neq 0), \\ -1 & (s=0). \end{cases}$$

[5] C'est-à-dire, la série de puissance $\Sigma(A_n+iB_n)z^n$ est convergente au point $z=e^{-2\pi i x}$ du cercle-unité.

[6] Voir [11], p. 269 pour $\Re(s)<0$, $0<u\leqq 1$. — Comme nous l'avons montré récemment, ceci est valable aussi pour $0\leqq\Re(s)<1$, $0<u<1$; cf. [8], p. 148.

[7] On sait que $\zeta(s,u)$ est engendré par la série $\sum_{m=0}^{\infty}(m+u)^{-s}$ $(\Re(s)>1, u>0)$ par prolongement analytique; $\zeta(s,1)=\zeta(s)$.

En observant que $\mathfrak{Z}_s(u)$ a été considérablement utilisée dans [9], nous nous référons aux faits suivants:[8] 1. $\mathfrak{Z}_s(u)$ est une fonction entière de la variable complexe s pour chaque u réel, fixé, et une fonction dérivable pour $u \in (-\infty, \infty), u \neq 0, \pm 1, \ldots,$ quelle que soit la valeur de s; 2. dans les extrémités de $0 \leqq u \leqq 1$, on a

(2.9) $$\mathfrak{Z}_s(1-0) = \Gamma(s)^{-1}\zeta(1-s),$$

(2.10) $$\begin{cases} \mathfrak{z}_s(+0) = \Gamma(s)^{-1}\zeta(1-s), \\ \mathfrak{z}_s(u) = \mathfrak{Z}_s(u) - \Gamma(s)^{-1}u^{s-1} \end{cases}$$

et il subsiste l'équation fonctionnelle

(2.11) $$\frac{\partial}{\partial u}\mathfrak{Z}_{s+1}(u) = \mathfrak{Z}_s(u) \qquad (u \neq 0, \pm 1, \ldots).$$

Cela posé, on voit immédiatement que la série (totalement divergente)

(2.12) $$2\sum_{n=1}^{\infty} \cos 2n\pi x$$

est sommable $(\mathfrak{W}_+)$ ou $(\mathfrak{W}_-)$ pour $x \neq 0, \pm 1, \ldots$ avec la somme -1; le même résultat s'obtient par la méthode $(C, 1)$.

§ 3. Application à la série de Fourier

1. Considérons la série de Fourier d'une fonction $f(u) \in L(0, 1)$:

(3.1) $$\begin{cases} \alpha_0 + 2\sum_{n=1}^{\infty}(\alpha_n \cos 2n\pi u + \beta_n \sin 2n\pi u), \\ \alpha_n = \int_0^1 f(t) \cos 2n\pi t\, dt, \quad \beta_n = \int_0^1 f(t) \sin 2n\pi t\, dt. \end{cases}$$

On sait maintenant que pour $0 < \theta < 1$

(3.2) $$\begin{cases} \sum_{\pm}(x, \theta) = \cos\frac{\pi\theta}{2}\sum_{n=1}^{\infty} 2(2n\pi)^{-\theta}(\alpha_n \cos 2n\pi x + \beta_n \sin 2n\pi x) \pm \\ \qquad \pm \sin\frac{\pi\theta}{2}\sum_{n=1}^{\infty} 2(2n\pi)^{-\theta}(\beta_n \cos 2n\pi x - \alpha_n \sin 2n\pi x) \end{cases}$$

converge en presque tout point x.[9]

[8] Cf. [9], p. 84; [8], p. 145—147.
[9] Cf. [9], p. 99.

2. Dans ce qui suit — pour assurer la convergence des séries (3.2) — nous *supposerons que* $f(u)$ *est mesurable et bornée dans* (0, 1) *et de période* 1, sauf si une autre condition est formulée; de plus, *soit toujours* $0 < x < 1$ *et* $0 < \theta < 1$.

THÉORÈME I. 1. *Pour que la série* (3.1) *soit sommable* ($\mathfrak{W}_+$) *en un point* x, *il faut et il suffit que la limite*

$$f\langle x+0\rangle = \lim_{\theta\to+0}\left(\theta\int_0^{\delta} f(x+t)t^{\theta-1}\,dt\right) \tag{3.3}$$

existe, $\delta > 0$ *désignant un nombre fixé arbitrairement petit.*[10]

Si (3.3) *existe pour une valeur* $\delta = \delta_1 > 0$, *elle est constante pour* $0 < \delta \leqq \delta_1$ *et fournit la somme* ($\mathfrak{W}_+$) *de* (3.1) *au point* x.

2. *En particulier,* $f\langle x+0\rangle = f(x+0)$ *pourvu que cette limite existe.*

3. *On a des énoncés entièrement analogues concernant la méthode* ($\mathfrak{W}_-$), *avec* $f(x-t)$ *pour* $f(x+t)$.

4. *Si* $f(u)$ *est continue en tous les points* x *d'un segment* $[a, b]$, (3.1) *est uniformément sommable* ($\mathfrak{W}_\pm$) *dans cet intervalle avec la somme* $f(x)$.

DÉMONSTRATION. 1. En utilisant une extension de la formule de Parseval (cf. [9], p. 99—100) ou la convergence "majorisée" des séries[11]

$$\sum_{n=1}^{\infty}\frac{\cos 2n\pi u}{(2n\pi)^{\theta}},\quad \sum_{n=1}^{\infty}\frac{\sin 2n\pi u}{(2n\pi)^{\theta}}$$

dans l'intervalle $0 < u < 1$ et le théorème de LEBESGUE sur l'intégration terme à terme, on trouve les séries (3.2) convergentes pour tout x ($\in (0, 1)$) et la formule (cf. (2.7), (2.8)).

$$\begin{cases} \sum_{n=1}^{\infty}(2n\pi)^{-\theta}\left[\alpha_n\cos\left(2n\pi x+\frac{\pi\theta}{2}\right)+\beta_n\sin\left(2n\pi x+\frac{\pi\theta}{2}\right)\right]= \\ \qquad =\int_0^1 f(u)\mathfrak{z}_\theta(u-x)\,du=\int_0^1 f(x+t)\mathfrak{z}_\theta(t)\,dt. \end{cases} \tag{3.4}$$

Soit x fixé et envisageons la dernière intégrale sous la forme suivante (cf. (2.10)):

$$\begin{cases} \int_0^1 f(x+t)\mathfrak{z}_\theta(t)\,dt=\int_0^1 f(x+t)\mathfrak{z}_\theta(t)\,dt+\Gamma(\theta)^{-1}\int_0^{\delta} f(x+t)t^{\theta-1}\,dt+ \\ \qquad +\Gamma(\theta)^{-1}\int_{\delta}^{1} f(x+t)t^{\theta-1}\,dt=I_1+I_2+I_3. \end{cases} \tag{3.5}$$

[10] Cf. le "principe de localisation" de [9], p. 95.

[11] Cf. par exemple [5], p. 33, 95.

Pour étudier le premier terme, nous observons

$$(3.6)\quad \begin{cases} 0 < \Gamma(\theta)^{-1}\zeta(1-\theta) - \mathfrak{z}_\theta(t) = \Gamma(\theta)^{-1}\sum_{n=1}^{\infty}[n^{\theta-1}-(n+t)^{\theta-1}] < \\ \quad < \Gamma(\theta)^{-1}\sum_{n=1}^{\infty}[n^{\theta-1}-(n+1)^{\theta-1}] = \Gamma(\theta)^{-1} \quad (0<t<1). \end{cases}$$

Comme $\Gamma(\theta)^{-1} \to +0 \quad (\theta \to +0)$, il en ressort que

$$(3.7)\quad \begin{cases} |I_1+\alpha_0| \leqq |I_1-\Gamma(\theta)^{-1}\zeta(1-\theta)\cdot\alpha_0| + |\alpha_0[\Gamma(\theta)^{-1}\zeta(1-\theta)+1]| \leqq \\ \quad \leqq \Gamma(\theta)^{-1}\int_0^1 |f(x+t)|dt + |\alpha_0|\,|\Gamma(\theta)^{-1}\zeta(1-\theta)+1| < \frac{\varepsilon}{3}, \end{cases}$$

quel que soit ε positif, à condition que θ soit choisi suffisamment petit: $\theta < \theta'_\varepsilon$.

Quant au deuxième terme, on peut établir

$$(3.8)\quad \begin{cases} \left|I_2-\theta\int_0^\delta f(x+t)\,t^{\theta-1}\,dt\right| = \theta|\Gamma(\theta+1)^{-1}-1|\,\left|\int_0^\delta f(x+t)\,t^{\theta-1}\,dt\right| \leqq \\ \quad \leqq K[1-\Gamma(\theta+1)^{-1}] < \frac{\varepsilon}{3} \end{cases}$$

avec $K = \sup\limits_{(0,1)}|f(u)|$, si $\theta < \theta''_\varepsilon(<1)$.

Enfin, pour $\delta \in (0, 1)$ fixé, il subsiste

$$(3.9)\quad |I_3| \leqq \Gamma(\theta)^{-1}\,K\log(1/\delta) < \frac{\varepsilon}{3},$$

étant supposé que $\theta < \theta'''_\varepsilon$.

Ainsi, en conclusion de (3. 5)—(3. 9), on obtient le résultat:

$$(3.10)\quad \left|\alpha_0+\int_0^1 f(x+t)\mathfrak{Z}_\theta(t)\,dt - \theta\int_0^\delta f(x+1)t^{\theta-1}\,dt\right| < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} = \varepsilon,$$

$\delta \in (0, 1)$ étant donné arbitrairement, pourvu que $0 < \theta < \theta^*_\varepsilon$; par (3. 4) l'inégalité (3. 10) fournit ce qu'il nous fallait montrer sur la sommation $(\mathfrak{W}_+)$ de (3. 1).

Remarquons que la différence

$$\theta\int_0^{\delta_1} f(x+t)t^{\theta-1}dt - \theta\int_0^\delta f(x+t)t^{\theta-1}dt \qquad (0<\delta<\delta_1)$$

tendant évidemment vers zéro pour $\theta \to +0$, l'existence de la limite du premier terme entraîne le même fait pour le second, avec l'identité des valeurs des limites.

2. Supposons que $f(u)$ possède une limite à droite en x. Alors on aura pour $0<\eta<\delta<1$, $M=\sup_{0\leqq t\leqq 1}|f(x+t)-f(x+0)|$:

$$\text{(3.11)}\quad \begin{cases} \left|\theta\int_0^\delta f(x+t)t^{\theta-1}\,dt-f(x+0)\right| \leqq \left|\theta\int_0^\delta [f(x+t)-f(x+0)]t^{\theta-1}\,dt\right|+ \\ \quad +|\delta^\theta-1|\,|f(x+0)| \leqq \sup_{(0,\eta)}|f(x+t)-f(x+0)|+\theta M\log(\delta/\eta) \end{cases}$$

et les derniers termes tendant respectivement vers 0 avec η et (pour η fixé) avec θ, leur somme devient aussi petite que l'on veut lorsque θ est choisi assez petit.

3. Dans le cas du procédé ($\mathfrak{W}_-$), nous n'avons à modifier que certains signes aux raisonnements précédents: le rôle de $f\langle x+0\rangle$ sera joué par

$$\text{(3.12)}\qquad f\langle x-0\rangle=\lim_{\theta\to+0}\left(\theta\int_0^\delta f(x-t)t^{\theta-1}\,dt\right).$$

4. Pour vérifier l'énoncé sur la sommabilité ($\mathfrak{W}_\pm$) *uniforme*, il suffit d'observer que 1° si $f(u)$ est continue dans $[a,b]$, on peut écrire (cf. (3.11)):

$$\text{(3.13)}\qquad \left|\theta\int_0^\delta f(x\pm t)t^{\theta-1}\,dt-f(x)\right|\leqq \underset{(a,b)}{\omega}(\eta;f)+\theta M^*\log(\delta/\eta)$$

pour $a\leqq x\leqq b$, avec $0<\eta<\delta<1$ et

$$M^*=\sup_{\substack{0\leqq u\leqq 1\\ a\leqq x\leqq b}}|f(u)-f(x)|,$$

$\underset{(a,b)}{\omega}(\eta;f)$ désignant le *module de continuité* sur $[a,b]$; 2° les bornes figurant devant les derniers termes de (3.7)—(3.9) sont indépendants de la variable x.

En effet, pour cette raison, on obtient en chaque point $x\in[a,b]$ simultanément

$$\left|\left(\alpha_0+\int_0^1 f(x+t)\mathfrak{z}_\theta(t)\,dt\right)-f(x)\right|<\varepsilon,$$

pourvu que $\theta<\theta_0=\theta_0(\varepsilon)$. C. q. f. d.

3. Ajoutons d'abord un *corollaire* du Théorème I: *la condition nécessaire et suffisante pour que* (3. 1) *ait* $f(x)$ *pour somme* ($\mathfrak{W}_\pm$) *en un point* x, *est que, pour un* $\delta\in(0,1)$ *convenable,*

$$\text{(3.14)}\qquad \langle f(u)-f(x)\rangle_{u=x\pm0}=0$$

avec la notation $\langle\varphi(u)\rangle_{u=x\pm0}=\varphi\langle x\pm0\rangle$ (cf. (3. 3), (3. 12)).

Nous mentionnons déjà maintenant que $f\langle x-0\rangle$, si elle existe, est en même temps égale à la "limite W_0" correspondante, c'est-à-dire

$$f\langle x-0\rangle = f_{[0]}(x). \tag{3.15}$$

(Cf. [9], p. 85—86 et (2.8).)

4. Il est aisé de voir qu'au cas de l'existence de $f(x+0)$ ou de $f(x-0)$, les résultats justement établis nous permettent de calculer simplement ces limites unilatérales de la série de Fourier (3.1) et ainsi ils donnent une solution du problème de Fejér, formulé dans l'introduction.

Citons un théorème de Fejér qui s'y rapporte (cf. [3], p. 166): si $f(u)$ est une fonction qui satisfait aux conditions de Dirichlet, on aura

$$\lim_{N\to\infty} \left\{ \alpha_0 + 2\sum_{n=1}^{N} \left[\alpha_n \cos n \left(2\pi x \pm \frac{\varrho}{N}\right) + \beta_n \sin n \left(2\pi x \pm \frac{\varrho}{N}\right)\right]\right\} = f(x\pm 0),$$

où ϱ désigne une racine positive de l'équation transcendante $\operatorname{Si}(\varrho) = \int_0^\infty t^{-1} \sin t\, dt = 0$. — Les considérations de [3] (en connection avec le phénomène de Gibbs) furent développées progressivement par Rogosinski [10].

§ 4. Sommabilité ($\mathfrak{W}_\pm$) et points de Lebesgue

1. En vertu de ce qui a été dit plus haut, les limites $f\langle x\pm 0\rangle$ peuvent être considerées respectivement comme les *généralisations* des limites ordinaires unilatérales $f(x\pm 0)$: voici la raison des notations. Appellons un point x *point* ($\mathfrak{W}_+$) ou *point* ($\mathfrak{W}_-$), suivant que $f\langle x+0\rangle$ ou $f\langle x-0\rangle$ existe; on voit alors que tout point *régulier* (au sens introduit par Lebesgue) est simultanément un point ($\mathfrak{W}_+$) et ($\mathfrak{W}_-$). Un tel point sera appelé brièvement *point* ($\mathfrak{W}$).

Or, on sait que la condition suffisante la plus générale utilisée dans la théorie de sommation des séries de Fourier est l'existence de

$$f_\circ(x) = \lim_{h\to 0}\left(\frac{1}{2h}\int_{x-h}^{x+h} f(t)\,dt\right), \tag{4.1}$$

c'est-à-dire que l'intégrale indéfinie $F(u)$ de $f(u)$ ait une dérivée symétrique au point x; notamment, (3.1) est sommable par une méthode $(C, r), r>1$ (et ainsi par celle d'Abel—Poisson) en x avec la somme $f_\circ(x)$, si cette limite

12 Cf. par exemple [14], p. 52—56; pour $r=2$ voir Lebesgue [6].

existe.[12] L'existence de $F'(x)$ ou la coexistence de

$$f_+(x)=F'_+(x)=\lim_{h\to+0}\left(h^{-1}\int_0^h f(x+t)dt\right) \tag{4.2}$$

et

$$f_-(x)=F'_-(x)=\lim_{h\to+0}\left(h^{-1}\int_0^h f(x-t)dt\right) \tag{4.3}$$

entraîne naturellement la condition (4. 1). — Il est convenable pour la suite de faire usage des notations *point* $(\mathfrak{L}_\circ)$, $(\mathfrak{L}_+)$ ou $(\mathfrak{L}_-)$, bien entendu un point où — respectivement — la limite (4. 1), (4. 2) ou (4. 3) existe; au lieu de "point $(\mathfrak{L}_+)$ et $(\mathfrak{L}_-)$ simultanément" on dira aussi *point* $(\mathfrak{L})$.

2. Précisons tout d'abord, que (cf. Théorème I) l'ensemble des points $(\mathfrak{W}_\pm)$ d'une fonction $f(u)$ bornée et mesurable est identique à celui des points x où (3.1) est sommable $(\mathfrak{W}_\pm)$, ainsi que: la série (3.1) est sommable (C, r), $r>1$ en chaque point $(\mathfrak{L}_\circ)$ ou $(\mathfrak{L})$ (dans ce dernier cas ayant la somme $\frac{1}{2}[f_+(x)+f_-(x)]$).

Cela posé, il subsiste le

Théorème II. 1. *Si x est un point $(\mathfrak{L}_+)$, il est aussi un point $(\mathfrak{W}_+)$ et on a $f_+(x)=f\langle x+0\rangle$; on peut substituer dans cet énoncé moins à plus.*

Donc, pour toute fonction $f(u)$ en question, l'ensemble des points $(\mathfrak{W})$ contient tous les points $(\mathfrak{L})$.

2. *$f(u)$ étant choisie convenablement, il existe un point $(\mathfrak{W})$, qui n'est pas un point $(\mathfrak{L}_+)$ ou $(\mathfrak{L}_-)$, ni un point $(\mathfrak{L}_\circ)$.*

Démonstration. 1. En posant $F_x(t)=\int_0^t f(x+u)du$, supposons que

$$\lim_{t\to+0}\frac{F_x(t)}{t}=f_+(x) \tag{4.4}$$

existe. Ainsi on aura

$$F_x(t)=f_+(x)t+\eta_x(t) \tag{4.5}$$

avec

$$\lim_{t\to+0}\eta_x(t)=0. \tag{4.6}$$

D'autre part, en intégrant par parties il vient $(0<\delta\leqq 1, 0<\theta<1)$

$$\int_0^\delta f(x+t)t^{\theta-1}dt=F_x(\delta)\delta^{\theta-1}+(1-\theta)\int_0^\delta F_x(t)t^{\theta-2}dt. \tag{4.7}$$

En utilisant (4.7) et (4.5), nous écrivons

$$(4.8)\qquad \begin{cases} \theta\int\limits_0^\delta f(x+t)t^{\theta-1}\,dt - f_+(x) = \\ = \theta\delta^{\theta-1}F_x(\delta) + f_+(x)[\delta^\theta(1-\theta)-1] + \theta(1-\theta)\int\limits_0^\delta \eta_x(t)t^{\theta-1}dt = \\ = J_1+J_2+J_3. \end{cases}$$

Si l'on observe les inégalités

$$(4.9)\qquad \begin{cases} |J_1| \leqq \theta\cdot\sup\limits_{(0,1)}(t^{-1}F_x(t)), \\ |J_2| < |f_+(x)|(\delta^\theta-1), \\ |J_3| < \sup\limits_{(0,\delta)}|\eta_x(t)|, \end{cases}$$

il s'ensuit de (4.8) que, pour δ convenablement fixé (cf. (4.6)) la différence en question devient inférieure en valeur absolue à tout nombre $\varepsilon > 0$ donné à l'avance, pourvu que θ soit suffisamment petit; autrement dit

$$(4.10)\qquad \lim_{\theta\to+0}\left(\theta\int\limits_0^\delta f(x+t)\,t^{\theta-1}dt\right)$$

existe et est égale à $f_+(x)$.

On montre d'une façon analogue, que l'existence de $f_-(x)$ entraîne celle de $f\langle x-0\rangle$ et l'identité des deux limites.

2. Discutons la fonction $g(x)$ (en escalier), définie dans l'intervalle $\left(0, \frac{1}{2}\right]$ par

$$(4.11)\qquad g(x) = \begin{cases} -1, \text{ si } 3^{-(\nu+1)} < x \leqq 2^{-1}3^{-\nu} & (\nu = 0, 1, \ldots), \\ +1, \text{ si } 2^{-1}3^{-\nu} < x \leqq 3^{-\nu} & (\nu = 1, 2, \ldots), \end{cases}$$

d'ailleurs paire et de période 1.

Il est clair que la totalité des discontinuités de $g(x)$ dans $\left[-\frac{1}{2}, \frac{1}{2}\right]$ se compose de l'origine et des points $\pm\frac{1}{3^\nu}$ $(\nu=1, 2, \ldots)$, $\pm\frac{1}{2}\cdot\frac{1}{3^\nu}$ $(\nu=1,2,\ldots)$; notamment, $x=0$ est un point de discontinuité de seconde espèce, le reste contient ceux de première espèce.

En même temps, considérons l'intégrale

$$(4.12)\qquad G(x) = \int\limits_{\frac{1}{2}}^x g(t)\,dt;$$

c'est une fonction impaire et partout continue, ayant aussi la période 1. $G(x)$ est représentée dans $\left(-\frac{1}{2}, \frac{1}{2}\right)$ par une ligne brisée (infinie), dont tous les sommets sont les points $\left(\frac{1}{2}, 0\right), \left(\frac{1}{3}, \frac{1}{6}\right), \left(\frac{1}{6}, 0\right), \left(\frac{1}{9}, \frac{1}{18}\right), \left(\frac{1}{18}, 0\right), \left(\frac{1}{27}, \frac{1}{54}\right)$, $\left(\frac{1}{54}, 0\right), \ldots$; on trouve $|G(x)| \leqq \frac{1}{2}|x|$ (avec égalité seulement pour $x = \pm 3^{-n}$, $n = 1, 2, \ldots$) de sorte que $G(0) = 0$.

Or, il est aisé de voir qu'aucune des limites (dérivées) $g_+(0) = G'_+(0)$, $g_-(0) = G'_-(0)$, $g_\circ(0) = G_{(1)}(0)$ ne peut être finie et déterminée, car le rapport $G(h)/h$ est oscillant également pour $h \to +0$ ou $h \to -0$; en conséquence, $x = 0$ n'est ni un point $(\mathfrak{L}_+)$ ou $(\mathfrak{L}_-)$, ni un point $(\mathfrak{L}_\circ)$.

Par contre, on obtient

$$\theta \int_0^{\frac{1}{2}} g(t) t^{\theta-1} dt = \theta \left(\int_{\frac{1}{8}}^{\frac{1}{2}} + \int_{\frac{1}{6}}^{\frac{1}{8}} + \int_{\frac{1}{9}}^{\frac{1}{6}} + \int_{\frac{1}{18}}^{\frac{1}{9}} + \cdots \right) =$$

$$= 2^{-\theta} - \left(\sum_{\nu=0}^{\infty} 2^{-\theta} 3^{-\nu\theta} - \sum_{n=1}^{\infty} 3^{-n\theta} \right) = 2^{-\theta} - 2 \frac{2^{-\theta} - 3^{-\theta}}{1 - 3^{-\theta}} \quad (0 < \theta < 1);$$

$$g\langle \pm 0 \rangle = \lim_{\theta \to +0} \left(\theta \int_0^{\frac{1}{2}} g(t) t^{\theta-1} dt \right) = 1 - 2 \frac{\log 3 - \log 2}{\log 3} = \frac{\log 4}{\log 3} - 1,$$

ce qui montre que l'origine est un point $(\mathfrak{W})$ de $g(x)$. C. q. f. d.

3. Remarquons, naturellement, qu'il n'y a aucune difficulté à construire — d'après le modèle de $g(x)$ — des fonctions avec un nombre prescrit de points $(\mathfrak{W})$, "exceptionnels" au sens donné plus haut. — Ces exemples prennent plus encore de l'importance en observant le fait (cf. un théorème de HARDY et LITTLEWOOD; [14], 10.45 (II)):[13] pour que la série de Fourier d'une fonction bornée soit sommable (C) en un point x, il faut et il suffit que x soit un point $(\mathfrak{L}_\circ)$ de la fonction. Après ce qui précède, *la série de Fourier de $g(x)$ n'est pas sommable (C, r) au point $x = 0$, quel que soit $r > 0$, par contre, elle est sommable $(\mathfrak{W}_\pm)$.*

Quant à la première partie de notre théorème, il est bien connu que l'on a pour toute fonction sommable $f(u)$ presque partout

$$\lim_{h \to 0} \left(h^{-1} \int_x^{x+h} |f(t) - f(x)| dt \right) = 0 \tag{4. 13}$$

[13] L'auteur est reconnaissant au professeur A. RÉNYI d'appeler son attention sur la proposition en question.

(LEBESGUE [6]). Si les points x avec (4.13) sont nommés *points* $|\mathfrak{L}|$, on verra qu'il s'agit là d'une catégorie spéciale des points ($\mathfrak{L}$), donc — pour $f(u)$ bornée — d'un sous-ensemble des points ($\mathfrak{W}$).

Il s'ensuit que (dans notre cas) *presque tout point x est un point* ($\mathfrak{W}$).

§ 5. Séries de Fourier dérivées; cas de séries de Dirichlet ordinaires

1. En dérivant p-fois la série (3.1), il résulte

$$(5.1)\qquad 2\sum_{n=1}^{\infty}(2n\pi)^p\left[\alpha_n\cos\left(2n\pi u+\frac{p\pi}{2}\right)+\beta_n\sin\left(2n\pi u+\frac{p\pi}{2}\right)\right]$$
$$(p=1,2,\ldots).$$

(5.1) est aussi une série de Fourier, à savoir celle de $f^{(p)}(u)$, si $f^{(p-1)}(u)$ est absolument continue dans $[0, 1]$.[14] Ainsi l'application du Théorème I fournit (entre autres) la proposition suivante:

Dans le cas où $f^{(p-1)}(u)$ $(p\geqq 1)$ est absolument continue et $f^{(p)}(u)$ bornée, pour la sommabilité ($\mathfrak{W}_\pm$) de (5.1) *en un point x il faut et il suffit que*

$$(5.2)\qquad f_\pm^{[p]}(x)=\lim_{\theta\to+0}\left(\theta\int_0^\delta f^{(p)}(x\pm t)\,t^{\theta-1}\,dt\right)$$

existe avec un $\delta>0$ fixé; si (5.2) *existe, alors on aura*

$$(5.3)\qquad \overset{(\mathfrak{W}_\pm)}{\sum_{n=1}^{\infty}}2(2n\pi)^p\left[\alpha_n\cos\left(2n\pi x+\frac{p\pi}{2}\right)+\beta_n\sin\left(2n\pi x+\frac{p\pi}{2}\right)\right]=f_\pm^{[p]}(x).$$ [15]

Observons que $f_\pm^{[0]}(x)=f\langle x\pm 0\rangle$ et, plus généralement, $f_\pm^{[p]}(x)$ peuvent être regardées comme extensions des dérivées ordinaires d'ordre p à droite et à gauche (de $f_+^{(p)}(x)$ et resp. de $f_-^{(p)}(x)$). En effet, si l'on écrit $f(u)$ pour $F(u)$ au Théorème II, la première fonction étant supposée absolument continue, on obtient que l'existence de $f_+^{[1]}(x)$, $f_-^{[1]}(x)$ découle de celle de $f'_+(x)$ resp. de $f'_-(x)$; d'autre part, $G(x)$ (cf. (4.12)) est une fonction particulière pour laquelle $G_\pm^{[1]}(0)$ existent simultanément, sans l'existence de $G'_\pm(0)$ (et de $G_{(1)}(0)$). L' énoncé complet s'obtient en considérant successivement $f'(u)$, $f''(u)$ etc.

[14] Celà implique, comme on le sait, l'existence de $f^{(p)}(u)$ p. p.

[15] M. ZAMANSKY [12] a traité le problème de trouver des méthodes linéaires sommant (5.1) partout à la dérivée symétrique $f_{(p)}(x)$ où cette dernière limite existe. — Voir aussi [13].

2. Quant à la sommabilité ($\mathfrak{W}_{\pm}$) de la série (naturellement généralisée)

$$(5.4)\qquad \sum_{n=1}^{\infty}\frac{2}{(2n\pi)^s}\left[\alpha_n\cos\left(2n\pi u-\frac{\pi s}{2}\right)+\beta_n\sin\left(2n\pi u-\frac{\pi s}{2}\right)\right],$$

où s désigne un nombre complexe, nous avons apparemment besoin de la convergence de

$$(5.5)\ \sum_{n=1}^{\infty}\frac{2}{(2n\pi)^{s+\theta}}\left[\alpha_n\cos\left(2n\pi u+\frac{\pi}{2}(-s\pm\theta)\right)+\beta_n\sin\left(2n\pi u+\frac{\pi}{2}(-s\pm\theta)\right)\right]$$
$$(0<\theta<1)$$

au point $u=x$ en question (cf. (2.3), (2.5)); pour assurer cela, nous supposerons que $f(u)$ est choisi dans une classe $L^q(0, 1)$ convenable (dépendante de s). En outre, nous pouvons établir aussi que le cas de la sommation ($\mathfrak{W}_-$) se distingue essentiellement au point de vue de la représentation de (5.5) sous la forme d'une intégrale.

Précisons:

Théorème III. 1. *Soit $f(u)\in L^q(0, 1)$ $(1\leqq q<\infty)$ et s_0 un nombre complexe avec $\Re(s_0)=1/q$.*

Pour que la série (5.4) soit sommable ($\mathfrak{W}_-$) avec $u=x$ et $s=s_0$, il faut et il suffit que

$$(5.6)\qquad \lim_{\theta\to+0}\int_0^{\delta}f(x-t)\,t^{s_0+\theta-1}\,dt\qquad(0<\delta\leqq 1)$$

existe; si cette condition est remplie, la somme ($\mathfrak{W}_-$) de (5.4) pour $u=x$, $s=s_0$ prend la valeur (cf. (2.10))

$$(5.7)\qquad \begin{cases} f_-^{[-s_0]}(x)=\displaystyle\int_0^1 f(x-t)\,\mathfrak{z}_{s_0}(t)\,dt+\\ \displaystyle +\Gamma(s_0)^{-1}\left[\int_{\delta}^{1}f(x-t)\,t^{s_0-1}\,dt+\lim_{\theta\to+0}\int_0^{\delta}f(x-t)\,t^{s_0+\theta-1}\,dt\right].\end{cases}$$

2. *Pourvu que $f^{(p-1)}(u)$ $(p\geqq 1$, entier$)$ soit absolument continue et $f^{(p)}(u)\in L^q(0,1)$ $(1\leqq q<\infty)$, la condition nécessaire et suffisante pour la sommabilité ($\mathfrak{W}_-$) de la série*

$$(5.8)\ \sum_{n=1}^{\infty}2(2n\pi)^{p-s_0}\left[\alpha_n\cos\left(2n\pi x+\frac{\pi}{2}(p-s_0)\right)+\beta_n\sin\left(2n\pi x+\frac{\pi}{2}(p-s_0)\right)\right]$$

se réduit à l'existence de la limite

$$(5.9)\qquad \lim_{\theta\to+0}\int_0^{\delta}f^{(p)}(x-t)\,t^{s_0+\theta-1}\,dt;$$

dans ce cas, il subsiste

$$(5.10)\quad \begin{cases} \overset{(\mathfrak{W}_-)}{\displaystyle\sum_{n=1}^{\infty}} 2(2n\pi)^{p-s_0}\left[\alpha_n \cos\left(2n\pi x+\frac{\pi}{2}(p-s_0)\right)+\right. \\ \left.+\beta_n \sin\left(2n\pi x+\frac{\pi}{2}(p-s_0)\right)\right]=\displaystyle\int_0^1 f^{(p)}(x-t)\mathfrak{z}_{s_0}(t)\,dt+ \\ +\Gamma(s_0)^{-1}\left[\displaystyle\int_\delta^1 f^{(p)}(x-t)t^{s_0-1}dt+\lim_{\theta\to+0}\int_0^\delta f^{(p)}(x-t)t^{s_0+\theta-1}dt\right]. \end{cases}$$

DÉMONSTRATION. 1. Dans notre hypothèse, on peut conclure du Théorème VII de [9], que

$$\sum_{n=1}^{\infty} 2(2n\pi)^{-(s_0+\theta)}\left[\alpha_n \cos\left(2n\pi x-\frac{\pi}{2}(s_0+\theta)\right)+\beta_n \sin\left(2n\pi x-\frac{\pi}{2}(s_0+\theta)\right)\right]=$$

$$=\cos\frac{\pi}{2}(s_0+\theta)\sum_{n=1}^{\infty} 2(2n\pi)^{-(s_0+\theta)}(\alpha_n \cos 2n\pi x+\beta_n \sin 2n\pi x)+$$

$$+\sin\frac{\pi}{2}(s_0+\theta)\sum_{n=1}^{\infty} 2(2n\pi)^{-(s_0+\theta)}(\alpha_n \sin 2n\pi x-\beta_n \cos 2n\pi x)$$

est convergente pour tout x et θ en question, sa somme étant donnée par l'intégrale

$$(5.11)\qquad \int_0^1 f(x-t)\mathfrak{Z}_{s_0+\theta}(t)\,dt.$$

Nous devons donc considérer (5.11) pour $\theta\to+0$.

En faisant usage de la décomposition $(0<\delta\leqq 1$, cf. (3.5))

$$(5.12)\quad \begin{cases} \displaystyle\int_0^1 f(x-t)\mathfrak{Z}_{s_0+\theta}(t)\,dt=\int_0^1 f(x-t)\mathfrak{z}_{s_0+\theta}(t)\,dt+ \\ +\Gamma(s_0+\theta)^{-1}\displaystyle\int_0^\delta f(x-t)\,t^{s_0+\theta-1}dt+\Gamma(s_0+\theta)^{-1}\int_\delta^1 f(x-t)t^{s_0+\theta-1}\,dt, \end{cases}$$

on voit que les limites de la première et troisième portions de notre intégrale s'obtiennent simplement par la substitution $\theta=0$, parce que — en vertu de [9], Th. III, 1 et IV — $\int_0^1 f(x-t)\mathfrak{z}_s(t)\,dt$, $\int_\delta^1 f(x-t)\,t^s dt$ sont des fonctions partout continues (de plus holomorphes) de la variable s. Puisque $\Gamma(s_0)\neq 0$, on en déduit immédiatement le critère (5.6) et puis, dans cette condition,

la formule de sommation:

$$(5.13)\qquad \overset{(\mathfrak{B}_-)}{\sum_{n=1}^{\infty}} \frac{2}{(2n\pi)^{s_0}}\left[\alpha_n \cos\left(2n\pi x-\frac{\pi s_0}{2}\right)+\beta_n \sin\left(2n\pi x-\frac{\pi s_0}{2}\right)\right]=f_-^{[-s_0]}(x).$$

2. Si $f^{(p-1)}(u)$ est absolument continue, alors

$$(5.14)\qquad \begin{cases} (2n\pi)^p\alpha_n=\alpha_n^{(p)}\cos\dfrac{p\pi}{2}-\beta_n^{(p)}\sin\dfrac{p\pi}{2}, \\ (2n\pi)^p\beta_n=\alpha_n^{(p)}\sin\dfrac{p\pi}{2}+\beta_n^{(p)}\cos\dfrac{p\pi}{2} \end{cases}$$

avec $\alpha_n^{(p)}=\int\limits_0^1 f^{(p)}(t)\cos 2n\pi t\,dt,\quad \beta_n^{(p)}=\int\limits_0^1 f^{(p)}(t)\sin 2n\pi t\,dt.$

Donc maintenant (5.8) s'écrit sous la forme

$$(5.15)\qquad \sum_{n=1}^{\infty} 2(2n\pi)^{-s_0}\left[\alpha_n^{(p)}\cos\left(2n\pi x-\frac{\pi s_0}{2}\right)+\beta_n^{(p)}\sin\left(2n\pi x-\frac{\pi s_0}{2}\right)\right]$$

et la seconde partie de l'énoncé découle de la première en l'appliquant pour $f^{(p)}(u)$ au lieu de $f(u)$.

3. En relation avec le Théorème III, 1 nous mentionnons que la partie correspondante du Théorème I en peut être obtenue pour s_0 réel, $q\to\infty$, si l'on identifie la classe $L^\infty(0,1)$ à la totalité des fonctions bornées, mesurables sur $(0,1)$. — On voit que les séries

$$(5.16)\qquad \begin{cases} \sum\limits_{n=1}^{\infty} 2(2n\pi)^{-s}(\alpha_n\cos 2n\pi x+\beta_n\sin 2n\pi x), \\ \sum\limits_{n=1}^{\infty} 2(2n\pi)^{-s}(\alpha_n\sin 2n\pi x+\beta_n\cos 2n\pi x) \end{cases}$$

convergentes pour $\Re(s)>q^{-1}$ (partie 1) et $\Re(s)>-p+q^{-1}$ (partie 2), respectivement, cessent de converger en général sur la frontière des demi-plans en question; apparemment, les résultats justement vérifiés nous permettent la sommation de (5.4) en certains points de $\Re(s)=q^{-1}$ ou $\Re(s)=-p+q^{-1}$ même dans des cas où ce n'est pas possible par d'autres méthodes usuelles.

Il est à noter que, par exemple, pour $f(u)$ paire on a la formule

$$(5.17)\qquad \begin{cases} \sum\limits_{n=1}^{\infty}\dfrac{\alpha_n\cos 2n\pi x}{n^s}=\dfrac{1}{4}(2\pi)^s\sec\dfrac{\pi s}{2}\left[\sum\limits_{n=1}^{\infty}\dfrac{2}{(2n\pi)^s}\alpha_n\cos\left(2n\pi x-\dfrac{\pi s}{2}\right)+\right. \\ \left.\qquad +\sum\limits_{n=1}^{\infty}\dfrac{2}{(2n\pi)^s}\alpha_n\cos\left(2n\pi x+\dfrac{\pi s}{2}\right)\right] \end{cases}$$

et les deux séries à droite (étant supposées convergentes) sont du type (5.4); donc le Théorème III est encore applicable à une classe assez large des séries de Dirichlet ordinaires.

4. Élucidons le rapport entre des dérivées généralisées introduites plus haut et les *"limites W_s"* au sens de [9]!

En premier lieu, par la définition de $f_{[s_0]}(x)$ (cf. [9], p. 85—86) il subsiste

$$(5.18)\quad \begin{cases} f_{[s_0]}(x)+\alpha_0 \mathfrak{Z}_{s_0}(x)= \\ \overset{(\mathfrak{W}_-)}{=} \sum_{n=1}^{\infty} 2(2n\pi)^{-s_0}\left[\alpha_n \cos\left(2n\pi x-\frac{\pi s_0}{2}\right)+\beta_n \sin\left(2n\pi x-\frac{\pi s_0}{2}\right)\right], \end{cases}$$

pourvu que les séries (5.16) convergent simultanément dans le demi-plan $\Re(s)>\Re(s_0)$. Ainsi, dans les conditions convenables sur $f(u)$, on aura (cf. (5.7))

$$(5.19)\qquad f_-^{[-s_0]}(x)=f_{[s_0]}(x)+\alpha_0 \mathfrak{Z}_{s_0}(x)$$

et, en particulier,

$$(5.20)\qquad f_-^{[p]}(x)=f_{[-p]}(x)\qquad (p=0, 1, \ldots)$$

(cf. [9], f. (3.3), (3.4)).

Il ressort que la somme $(\mathfrak{W}_-)$ de la série (5.17) peut s'exprimer simplement au moyen de $f_{[s]}(x)$ et la représentation obtenue fournit en même temps le prolongement analytique (relativement à s) de la somme.

5. Enfin nous enregistrons — cela s'impose — la possibilité de généraliser la méthode $(\mathfrak{W}_\pm)$ de manière que la sommabilité (C) ou (A) de (2.3) (au lieu de la convergence de cette série) soit supposée pour $0<\theta<1$. En tout cas, il semble que la variante de la définition

$$(5.21)\quad \begin{cases} \overset{(\mathfrak{W}_\pm^*)}{}\left[A_0+2\sum_{n=1}^{\infty}(A_n \cos 2n\pi x+B_n \sin 2n\pi x)\right]= \\ =A_0+\lim_{\theta\to+0}\sum_{n=1}^{\infty}\frac{2e^{-n\theta}}{(2n\pi)^\theta}\left[A_n \cos\left(2n\pi x\pm\frac{\pi\theta}{2}\right)+B_n \sin\left(2n\pi x\pm\frac{\pi\theta}{2}\right)\right] \end{cases}$$

soit bien utilisable.

(Reçu le 7 décembre 1959.)

Ouvrages cités

[1] L. Fejér, Sur les fonctions bornées et intégrables, *Comptes Rendus Acad. Sci. Paris*, **131** (1900), p. 984—987.

[2] L. Fejér, Untersuchungen über Fouriersche Reihen, *Math. Annalen*, **58** (1904), p. 51—69.

[3] L. Fejér, Über die Bestimmung des Sprunges der Funktion aus ihrer Fourierreihe, *Journal f. reine u. angew. Math.*, **142** (1913), p. 165—188.

[4] G. H. Hardy and M. Riesz, *The general theory of Dirichlet's series* (Cambridge, 1944).

[5] G. H. Hardy and W. W. Rogosinski, *Fourier series* (Cambridge, 1944).

[6] H. Lebesgue, Recherches sur la convergence des séries de Fourier, *Math. Annalen*, **61** (1905), p. 251—280.

[7] M. Lecat, *Bibliographie des séries trigonométriques* (Bruxelles—Louvain, 1921).

[8] M. Mikolás, Mellinsche Transformation und Orthogonalität bei $\zeta(s,u)$; Verallgemeinerung der Riemannschen Funktionalgleichung von $\zeta(s)$, *Acta Sci. Math. Szeged*, **17** (1956), p. 143—164.

[9] M. Mikolás, Differentation and integration of complex order of functions represented by trigonometrical series and generalized zeta-functions, *Acta Math. Acad. Sci. Hung.*, **10** (1959), p. 77—124.

[10] W. Rogosinski, Reihensummierung durch Abschnitts-Koppelungen, *Math. Zeitschrift*, **25** (1926), p. 132—149.

[11] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, 4. edition (Cambridge, 1952).

[12] M. Zamansky, Sur la sommation des séries de Fourier dérivées, *Comptes Rendus Acad. Sci. Paris*, **231** (1950), p. 1128—1120.

[13] M. Zamansky, *La sommation des séries divergentes* (Paris, 1954).

[14] A. Zygmund, *Trigonometrical series*, 2. edition (New York, 1952).

A CHARACTERIZATION OF TORSION ABELIAN GROUPS ONCE BASIC SUBGROUPS HAVE BEEN CHOSEN[1]

By

D. K. HARRISON (Haverford, Pa, USA)

(*Presented by* L. RÉDEI)

An abelian group which has every element of finite order is called a torsion group and it remains one of the main problems in the study of infinite abelian groups to characterize all torsion groups. If every element has order a power of a fixed prime p, then the group is called p-primary, and it is simple and well known that every torsion group is uniquely a direct sum of p-primary groups (and, of course, the converse is true also). If G is an abelian group and n an integer, $n \cdot G$ denotes the group of elements of the form $n \cdot x$ for some x in G. G is called divisible if $n \cdot G = G$ for all integers n. If Q is the additive group of rational numbers and Z is the subgroup of integers, then the set $Z(p^\infty)$ of all elements of Q/Z of p-power order is an example of a divisible p-primary group. In fact, the direct sums of copies of $Z(p^\infty)$ give precisely all of the divisible p-primary groups (for the theory of divisible groups see Chapter 3 of [1], for instance). Moreover, every p-primary group is a direct sum of a p-primary divisible group and a p-primary reduced group, where a reduced group is one that has no divisible subgroups other than 0. Hence the study of torsion groups is reduced to the study of p-primary reduced groups.

A subgroup H of a group G is called a pure subgroup if $(n \cdot G) \cap H = n \cdot H$ for all integers n. If G is a p-primary reduced group, H is called a basic subgroup of G if: 1. H is a pure subgroup of G, 2. H is a direct sum of cyclic groups (all of which will be of p-power order), and 3. G/H is divisible (and thus will be isomorphic to a direct sum of copies of $Z(p^\infty)$) Hence, if G has a basic subgroup, it is made up of a direct sum of copie. of $Z(p^\infty)$ "put in a pure fashion on top of" a direct sum of cyclic groupss Now it turns out that every reduced p-primary group G has a basic subgroup, and that any two basic subgroups of G are isomorphic (see Chapter 5 of [1]). Hence to study all p-primary reduced groups we may take a fixed. arbitrary direct sum S of cyclic groups of p-power orders and find all

1 This work was done in part while the author was on a Faculty Summer Study Grant from the Danforth Foundation.

p-primary reduced groups which have a basic subgroup isomorphic to S. The difficulty in this procedure is that a general p-primary reduced group has many basic subgroups, all of which are isomorphic of course, but none of which is more natural than the others. It is the purpose of this note to factor out this difficulty by studying a p-primary group together with a chosen basic subgroup, rather than studying just the reduced group itself. This will throw a burden of responsibility for a lack of classification of torsion groups on a lack of understanding of what the different basic subgroups of a given p-primary reduced group are.

Once and for all let p be a fixed prime number and let S be a fixed direct sum of cyclic groups of p-power orders. By a group pair (f, G) we shall mean a p-primary reduced group G together with an isomorphism f of S into G with $f(S)$ pure in G and $G/f(S)$ divisible (i. e. f is an isomorphism of S onto a basic subgroup of G). This will be our object of study, and this is what we mean by a p-primary reduced group together with a chosen basic subgroup. Two group pairs (f_1, G_1) and (f_2, G_2) will be called equivalent if there exists an isomorphism g of G_1 onto G_2 with $f_2 = g \cdot f_1$. We will proceed to characterize all group pairs (up to equivalence) as follows:

Let $S = \Sigma S(m)$ where the sum is direct and over all positive integers m and where $S(m)$ is a direct sum of α_m cyclic groups of order p^m for each m. Let S^* be the torsion subgroup (i. e. subgroup of elements of finite order) of $\Pi S(m)$ where the product (i. e. complete direct sum) is over all positive integers m. It can be shown that S^* is the torsion subgroup of the completion of S with respect to the $p^m \cdot S$ as spheres about the identity. At any rate the group S^*/S has a particularly simple structure, being in fact a direct sum of $\beta = 2^{\alpha \cdot \aleph_0}$ copies (or zero copies if $\alpha = 0$) of $Z(p^\infty)$ where α is the smallest cardinal which is greater than or equal to all but a finite number of the α_m. Now if we let $I_S = \mathrm{Hom}\,(Z(p^\infty), S^*/S) = \mathrm{Hom}\,(Z(p^\infty), \sum_\beta Z(p^\infty))$ (Hom (A, B) represents the group of all homomorphisms from A to B), then I_S also has a particularly simple structure. It is a direct product of $\alpha \cdot \aleph_0$ copies of the p-adic integers and is in fact the only torsion-free group (a group is torsion-free if zero is the only element of finite order) I which is complete with respect to taking the $p^m \cdot I$ as spheres about the identity and which has β as the dimension of $I/p \cdot I$ as vector space over the prime field of characteristic p. This metric topology furnished by taking the $p^m \cdot I_S$ as spheres about the identity is crucial. A subgroup H of I_S will be closed if and only if I_S/H is reduced. The purpose of this note is to establish a natural one-one correspondence between the equivalence classes of group pairs and the closed subgroups of I_S. We proceed now to proofs.

We will require a knowledge of the first and second sections of [2]. In particular, the concept of a co-torsion group is fundamental to our approach. If we write E for Ext $(Q/Z, S)$, by the second section of [2], S can be thought of as the torsion subgroup of E. We write E_w for $\cap n \cdot E$ where the intersection is over all positive integers n. But since $S \cap E_w \subset \cap n \cdot S = 0$, E_w is torsion-free. Since E/E_w is clearly reduced, Hom $(Q, E/E_w) = 0$. Hence we have the exact sequence

$$0 = \mathrm{Hom}\,(Q, E/E_w) \to \mathrm{Ext}\,(Q, E_w) \to \mathrm{Ext}\,(Q, E) = 0$$

where Ext $(Q, E) = 0$ by Section 2 of [2]. Thus Ext $(Q, E_w) = 0$ and E_w is a torsion-free co-torsion group (it is reduced since E is reduced by Section 2 of [2]). To calculate the torsion group which E_w corresponds to we consider the exact sequence

$$\mathrm{Tor}\,(Q/Z, E_w) = 0 \to \mathrm{Tor}(Q/Z, E) \to \mathrm{Tor}\,(Q/Z, E/E_w) \to (Q/Z) \otimes E_w \to (Q/Z) \otimes E.$$

We recall that Tor $(Q/Z, G)$ is isomorphic to the torsion subgroup G_t of G. Since E is adjusted, E/E_t is divisible and thus $(Q/Z) \otimes (E/E_t) = 0$ since Q/Z is a torsion group. Also $(Q/Z) \otimes E_t = 0$ and thus the sequence

$$(Q/Z) \otimes E_t \to (Q/Z) \otimes E \to (Q/Z) \otimes (E/E_t)$$

gives that $(Q/Z) \otimes E = 0$. Thus the earlier sequence gives

$$0 \to S \to (E/E_w)_t \to (Q/Z) \otimes E_w \to 0.$$

But this sequence is naturally isomorphic term by term to

$$0 \to S \to S^* \to S^*/S \to 0,$$

since E/E_w is complete in its natural metric by the last part of Section 3 of [2], and all completions of S are naturally isomorphic. This means by Section 2 of [2] that Hom $(Q/Z, S^*/S)$ is naturally isomorphic to E_w. That $S^*/S \simeq \sum_\beta Z(p^\infty)$ where $\beta = 2^{\alpha \cdot \aleph_0}$ (or, zero if $\alpha = 0$) can be shown using Theorem 33.4 of [1]. Hence E_w and a direct product of $\alpha \cdot \aleph_0$ copies of the p-adic integers are torsion-free co-torsion groups with the same invariants (see Section 2 of [2]) and thus are isomorphic.

We must now set up a correspondence between equivalence classes of group pairs and closed subgroups of E_w. We know the structure of E_w, so let I be any closed subgroup of E_w. By Lemma 3. 5 of [2] any subgroup H of E_w will be closed if and only if E_w/H is reduced. Hence E_w/I is reduced. We let G_I be the torsion subgroup of E/I and f_I be the homomorphism from S to G_I where $f_I(s) = s + I$ for all $s \in S$. Then (f_I, G_I) is easily seen to be a group pair.

If (f, G) is a group pair, we have the pure exact sequence $0 \to S \xrightarrow{f} G \to \to G/f(S) \to 0$ which gives rise to the sequence

$$\operatorname{Hom}(Q/Z, G) = 0 \to \operatorname{Hom}(Q/Z, G/f(S)) \xrightarrow{\delta} \operatorname{Ext}(Q/Z, S) \to \operatorname{Ext}(Q/Z, G) \to 0 = \\ = \operatorname{Ext}(Q/Z, G/f(S)).$$

We denote the image of δ in $E = \operatorname{Ext}(Q/Z, S)$ by $I_{(f,G)}$. Since the original sequence was pure, $I_{(f,G)}$ is actually in $P\operatorname{Ext}(Q/Z, S)$ which by Corollary 5.3 of [3] is E_w. Now by the above sequence, $(E/I_{(f,G)}) \simeq \operatorname{Ext}(Q/Z, G)$ which by Section 2 of [2] is reduced and has its torsion subgroup naturally isomorphic to G. By Lemma 3.5 of [2], $I_{(f,G)}$ is closed in E, and by an examination of the maps involved (f_I, G_I) is equivalent to (f, G) where $I = I_{(f,G)}$. It is clear that equivalent group pairs give the same closed subgroup of E_w.

It only remains to show that if I is a closed subgroup of $E_w, f = f_I$, and $G = G_I$, then $I_{(f,G)} = I$. The exact sequences $\operatorname{Ext}(Q, E) \to \operatorname{Ext}(Q, E/I) \to 0$ and $(Q/Z) \otimes E \to (Q/Z) \otimes (E/I) \to 0$ give that E/I is an adjusted co-torsion group. Hence we have by the theory of co-torsion groups that

$$0 \to I \to E \to E/I \to 0$$

implies

$$\begin{array}{ccccccccc}
0 & \to & \operatorname{Tor}(Q/Z, E) & \to & \operatorname{Tor}(Q/Z, E/I) & \to & I \otimes (Q/Z) & \to & 0 \\
 & & \wr & & \wr & & & & \\
0 & \longrightarrow & S & \xrightarrow{\quad f \quad} & G & & & &
\end{array}$$

which implies

$$\begin{array}{ccccccccc}
0 & \to & \operatorname{Hom}(Q/Z, I \otimes (Q/Z)) & \to & \operatorname{Ext}(Q/Z, S) & \to & \operatorname{Ext}(Q/Z, G) & \to & 0 \\
 & & \wr & & \wr & & \wr & & \\
0 & \longrightarrow & I & \longrightarrow & E & \longrightarrow & E/I & \to & 0.
\end{array}$$

This proves our result.

Since our duality is natural, all the natural properties of closed subgroups of E_w must correspond to natural properties of group pairs. We write $(f_1, G_1) \leqq \leqq (f_2, G_2)$ if there exists a homomorphism g of G_1 to G_2 with $g \cdot f_1 = f_2$. Then it can be shown that g, if it exists, is unique, and that $(f_1, G_1) \leqq (f_2, G_2)$ if and only if $I_1 \subset I_2$ where $I_1 = I_{(f_1, G_1)}$ and $I_2 = I_{(f_2, G_2)}$. The kernel of g will be isomorphic to the torsion subgroup of the adjusted part (see Section 2 of [2]) of I_2/I_1, and the cokernel of g (G_2 modulo the image of g) will be the divisible group dual (in the sense of [2]) to the torsion-free part of I_2/I_1. Clearly, (ident, S) will be less than or equal to all group pairs, and (inj, S^*) will be greater than or equal to all group pairs. G of (f, G) will have no elements of infinite height if and only if $I_{(f,G)}$ is a direct summand of E_w (since this is equivalent to $E_w/I_{(f,G)}$ being torsion-free). The pair $(f, G/G_w)$ (where $G_w = \bigcap n \cdot G$)

will correspond to the subgroup I which makes $I/I_{(f,G)}$ the adjusted part of $E_w/I_{(f,G)}$. The proofs of these facts are not at all immediate but are straightforward.

If (f^1, G^1) and (f^2, G^2) are group pairs with respect to S^1 and S^2, respectively, then $(f^1 \oplus f^2, S^1 \oplus S^2)$ is a group pair with respect to $S^1 \oplus S^2$. Since $\mathrm{Ext}\,(Q/Z, S^1 \oplus S^2) \simeq \mathrm{Ext}\,(Q/Z, S^1) \oplus \mathrm{Ext}\,(Q/Z, S^2)$, our duality can be applied to this situation.

The question of finding conditions on two closed subgroups I and I' of E_w equivalent to $G_I \simeq G_{I'}$ (this would characterize all torsion abelian groups) seems prohibitively difficult. Rather it is hoped that group pairs can be used in problems where there is some freedom in choosing basic subgroups.

It seems worth remarking that our treatment of group pairs gives something new essentially only in the case of groups with elements of infinite height, since (f, G) with $G_w = 0$, can be studied by considering the corresponding subgroup of S^*/S which exists because $(f, G) \leqq (\mathrm{inj}, S^*)$. In fact, what this note has really done is to allow us to consider any group pairs with the same ease as we can consider those whose group has no elements of infinite height.

HAVERFORD COLLEGE, HAVERFORD, PENNSYLVANIA, USA
AND
UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA, PENNSYLVANIA, USA

(Received 11 December 1959)

Bibliography

[1] L. Fuchs, *Abelian groups* (Budapest, 1958).

[2] D. K. Harrison, Infinite abelian groups and homological methods, *Annals of Math.*, **69** (1959), pp. 366—391.

[3] R. J. Nunke, Modules of extentions over Dedekind rings, *Illinois Journal of Math.*, **3** (1959), pp. 222—241.

ÜBER EINE METHODE ZUR NUMERISCHEN LÖSUNG DER POISSONSCHEN DIFFERENZENGLEICHUNG FÜR BELIEBIGE GEBIETE

Von

E. EGERVÁRY (Budapest), Mitglied der Akademie[1]

Problemstellung

Eine vergleichende Betrachtung der modernen Hilfsmittel der angewandten Mathematik bekräftigt besonders die an und für sich evidente Feststellung, daß es keine unveränderlich gültigen matematischen Modelle der Empirie gibt. Während aber die diskreten (finiten) Modelle sich auch schon früher mit den kontinuierlichen (infinitesimalen) Modellen prinzipiell gleichberechtigt erwiesen haben, weisen die numerischen Lösungsmethoden für diskrete Modelle nur seit dem Einsatz der programmgesteuerten, automatischen Rechenmaschinen eine entsprechende Entwicklung auf.

Die finiten Analoga der linearen partiellen Differentialgleichungen sind bekanntlich die Differenzengleichungen, d. h. ein spezielles System von linearen algebraischen Gleichungen. Man kann jedoch feststellen, daß — obwohl die reine Theorie der linearen algebraischen Gleichungen als längst abgeschlossen betrachtet werden kann — die bisher bekannten numerischen Lösungsmethoden noch nicht allen Anforderungen der praktischen Brauchbarkeit genügen.

Im besonderen wächst die Anzahl der auszuführenden arithmetischen Operationen bei zunehmender Anzahl der Unbekannten in vielen Aufgaben der Physik und Technik derartig schnell, daß manchmal sogar die Vorteile der elektronischen Rechenmaschinen illusorisch werden.

Für die Poissonschen (und ähnlichen) Differenzengleichungen sind in der neueren Literatur mehrere Lösungsmethoden angegeben worden, die meisten dieser Methoden sind jedoch nur auf ein rechteckiges Gebiet anwendbar.

Wir wollen in dieser Arbeit für die Poissonsche Differenzengleichung bezüglich eines beliebigen Gebietes eine numerische Lösungsmethode entwickeln, welche folgendermaßen charakterisiert werden kann:

1. Die Laplacesche Operatormatrix (welche auch in der Poissonschen Differenzengleichung vorkommt) kann in Blöcke partitioniert werden, welche vertauschbar sind. Für die Invertierung solcher Matrizen ist vom Verfasser ein Algorithmus entwickelt worden, welche die Ordnung der zu invertierenden Matrizen, also auch die Anzahl der auszuführenden Rechenoperationen wesentlich herabsetzt [2].

[1] Aus dem Nachlaß des Verfassers, bearbeitet durch P. Rózsa.

2. Ein beliebiges, aus Gitterpunkten bestehendes Gebiet kann immer in ein Rechteckgebiet eingebettet werden, und aus der (als bekannt vorausgesetzten) Inversen dieses Rechteckgebietes kann man die zum beliebigen Gebiete gehörige Inverse durch besonders einfache „rangvermindernde" Operationen berechnen [3].

Ähnliche Lösungsmethoden kann man auch für die biharmonischen und anderen linearen Differenzengleichungen entwickeln.

Man wird sehen, daß die in unserer Methode vorkommenden Rechenoperationen in hohem Grade einförmig sind, also eine bequeme Programmierung für automatische Rechenmaschinen gestatten.

I. Die Matrix des zweidimensionalen Laplaceschen Operators

In einem quadratischen Netz von Gitterpunkten

	x_{01}	x_{02}	x_{0j}	x_{0m}	
x_{10}	x_{11}	x_{12}	x_{1j}	x_{1m}	$x_{1,m+1}$
x_{20}	x_{21}	x_{22}	x_{2j}	x_{2m}	$x_{2,m+1}$
x_{i0}	x_{i1}	x_{i2}	x_{ij}	x_{im}	$x_{i,m+1}$
x_{n0}	x_{n1}	x_{n2}	x_{nj}	x_{nm}	$x_{n,m+1}$
	$x_{n+1,1}$	$x_{n+1,2}$	$x_{n+1,j}$	$x_{n+1,m}$	

wird der Laplacesche Differenzenoperator im Punkte (i,j) durch den Ausdruck

$$x_{i,j-1}-2x_{ij}+x_{i,j+1}+x_{i+1,j}-2x_{ij}+x_{i-1,j}=$$
$$=-(4x_{ij}-x_{i,j-1}-x_{i,j+1}-x_{i-1,j}-x_{i+1,j})$$

gegeben.

Werden bei Zugrundelegung eines Rechteckgebietes die nm Unbekannten x_{ij} $(i=1,2,\ldots,n;\ j=1,2,\ldots,m)$ in der lexikographischen Anordnung

$$x_{11}, x_{12}, \ldots, x_{1m}, x_{21}, x_{22}, \ldots, x_{2m}, \ldots, x_{n1}, x_{n2}, \ldots, x_{nm}$$

geschrieben, so lautet das Poissonsche Gleichungssystem bei Berücksichtigung der verschwindenden Randwerte $x_{0j}=x_{n+1,j}=x_{i0}=x_{i,m+1}=0$:

$$\begin{array}{lllll}
4x_{11}-x_{12} & -x_{21} & & & =p_{11}, \\
-x_{11}+4x_{12}-x_{13} & & -x_{22} & & =p_{12}, \\
 & & & & \ \vdots \\
 & & -x_{n-1,m} & -x_{n,m-1}+4x_{nm} & =p_{nm}.
\end{array}$$

Die Laplace—Poissonsche Koeffizientenmatrix ist also für ein Rechteckgebiet bei verschwindenden Randwerten

$$\mathbf{L}=\left[\begin{array}{c:c:c:c} \begin{matrix} 4 & -1 & & \\ -1 & 4 & & \\ & & \ddots & \\ & & -1 & 4 \end{matrix} & \begin{matrix} -1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{matrix} & & \\ \hdashline \begin{matrix} -1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{matrix} & \begin{matrix} 4 & -1 & & \\ -1 & 4 & & \\ & & \ddots & \\ & & & 4 \end{matrix} & & \\ \hdashline & & \ddots & \\ \hdashline & & & \begin{matrix} 4 & -1 & & \\ -1 & 4 & & \\ & & \ddots & \\ & & -1 & 4 \end{matrix} \end{array}\right]=$$

$$=\left[\begin{array}{c:c:c:c} \mathbf{K} & -\mathbf{E} & \cdots & \mathbf{O} \\ \hdashline -\mathbf{E} & \mathbf{K} & \cdots & \mathbf{O} \\ \hdashline \cdot & \cdot & \cdots & \\ \hdashline \mathbf{O} & \cdot & & \mathbf{K} \end{array}\right].$$

Diese Matrix läßt sich also in n^2 Blöcke m-ter Ordnung partitionieren, welche paarweise vertauschbar sind. Im Abschnitt II werden wir für solche Matrizen ein vereinfachtes Inversionsverfahren angeben, welche ausschließlich aus rationalen Operationen besteht.

Wir wollen hier noch eine andere Darstellung von **L** einführen, welche die Inversion mit Hilfe der Spektralzerlegung ermöglichen wird. Wir definieren zunächst (in unmittelbarem Anschluß an den Begriff des direkten Produktes) das „direkte Polynom" von zwei quadratischen Matrizen **A** und **B** beliebiger Ordnung folgendermaßen:

$$\varphi(\mathbf{A}, \mathbf{B})=\sum_p \sum_q c_{pq} \mathbf{A}^p \cdot \times \mathbf{B}^q .$$

Schreiben wir nun **L** in der Form

$$
\mathbf{L}=\left[\begin{array}{cccc|cccc|c|cccc}
2 & -1 & & & & & & & & & & & \\
-1 & 2 & & & & & & & & & & & \\
 & & \ddots & & & & & & & & & & \\
 & & & 2 & & & & & & & & & \\
\hline
 & & & & 2 & -1 & & & & & & & \\
 & & & & -1 & 2 & & & & & & & \\
 & & & & & & \ddots & & & & & & \\
 & & & & & & & 2 & & & & & \\
\hline
 & & & & & & & & \ddots & & & & \\
\hline
 & & & & & & & & & 2 & -1 & & \\
 & & & & & & & & & -1 & 2 & & \\
 & & & & & & & & & & & \ddots & \\
 & & & & & & & & & & & & 2
\end{array}\right]+
$$

$$
+\left[\begin{array}{cccc|cccc|c|cccc}
2 & & & & -1 & & & & & & & & \\
 & 2 & & & & -1 & & & & & & & \\
 & & \ddots & & & & \ddots & & & & & & \\
 & & & 2 & & & & -1 & & & & & \\
\hline
-1 & & & & 2 & & & & & & & & \\
 & -1 & & & & 2 & & & & & & & \\
 & & \ddots & & & & \ddots & & & & & & \\
 & & & -1 & & & & 2 & & & & & \\
\hline
 & & & & & & & & \ddots & & & & \\
\hline
 & & & & & & & & & 2 & & & \\
 & & & & & & & & & & 2 & & \\
 & & & & & & & & & & & \ddots & \\
 & & & & & & & & & & & & 2
\end{array}\right],
$$

so erkennt man unmittelbar, daß $\mathbf{L}$ sich als direktes Polynom der Kontinuanten $\mathbf{C}_m$ und $\mathbf{C}_n$,

$$\mathbf{C}_k = \begin{array}{c} \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & & \\ & & & \ddots & \\ & & & -1 & 2 \end{bmatrix} \end{array} \begin{array}{l} (1 \\ (2 \\ (3 \\ \\ (k \end{array}, \quad \mathbf{E}_k = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix},$$

in folgender Form ausdrücken läßt:

$$\mathbf{L} = \mathbf{C}_n \cdot \times \mathbf{E}_m + \mathbf{E}_n \cdot \times \mathbf{C}_m .$$

II. Inversion der Laplaceschen Operatormatrix für ein Rechteckgebiet

Erste Methode

Es sei $\mathbf{A}$ eine quadratische nichtsinguläre Matrix nm-ter Ordnung, welche als Hypermatrix

$$[\mathbf{A}_{ij}] \qquad (i, j = 1, 2, \ldots, n)$$

aus n^2 paarweise vertauschbaren Blöcken besteht, d. h.

$$\mathbf{A}_{ij}\mathbf{A}_{kl} = \mathbf{A}_{kl}\mathbf{A}_{ij} \qquad (i, j, k, l = 1, 2, \ldots, n).$$

Die quadratische Matrix m-ter Ordnung

$$\sum_{\nu} \pm \mathbf{A}_{1\nu_1}\mathbf{A}_{2\nu_2}\ldots\mathbf{A}_{n\nu_n},$$

welche (wegen der Vertauschbarkeit der Blöcke) wohldefiniert ist, soll die *Determinantenmatrix* der Hypermatrix $[\mathbf{A}_{ij}]$ genannt und mit $\mathbf{Det}\,[\mathbf{A}_{ij}]$ bezeichnet werden. In Verallgemeinerung einer von I. SCHUR stammenden Formel kann man leicht zeigen, daß die gewöhnliche Determinante $\mathrm{Det}\,\mathbf{Det}\,[\mathbf{A}_{ij}]$ der Determinantenmatrix m-ter Ordnung der gewöhnlichen Determinante $\mathrm{Det}\,\mathbf{A}$ der ursprünglichen Matrix $\mathbf{A}$ nm-ter Ordnung gleich ist.

Wenn also die Matrix nm-ter Ordnung invertierbar ist, so ist auch die Determinantenmatrix $\mathbf{Det}\,[\mathbf{A}_{ij}]$ invertierbar.

Aus der Matrix $(n-1)m$-ter Ordnung, welche aus $\mathbf{A}$ durch Tilgung der i-ten Blockzeile und der j-ten Blockspalte entsteht (welche also aus $(n-1)^2$ vertauschbaren Blöcken m-ter Ordnung besteht), kann nach der obigen Definition die Determinantenmatrix m-ter Ordnung gebildet werden. Diese Matrix, multipliziert mit $(-1)^{i+j}$, soll die zum Blocke $\mathbf{A}_{ij}$ gehörige algebraische *Komplementmatrix* genannt und mit $\mathbf{A}^*_{ji}$ bezeichnet werden.

Die Rolle der gewöhnlichen adjungierten Matrix wird nun von der Hypermatrix

$$[\mathbf{A}^*_{ij}] \qquad (i, j = 1, 2, \ldots, n)$$

übernommen. Werden nämlich in die skalare Identität, welche die Elemente einer gewöhnlichen Matrix und die Elemente ihrer Adjungierten verbindet, die vertauschbaren Blöcke $\mathbf{A}_{ij}$ von $\mathbf{A}$, sowie die (mit diesen Blöcken und untereinander vertauschbaren) Komplementmatrizen $\mathbf{A}^*_{ij}$ substituiert, so erhalten wir die für unsere Zwecke grundlegende Identität

$$\begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ \cdot & \cdots & \cdot \\ \mathbf{A}_{n1} & \cdots & \mathbf{A}_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{A}^*_{11} & \cdots & \mathbf{A}^*_{1n} \\ \cdot & \cdots & \cdot \\ \mathbf{A}^*_{n1} & \cdots & \mathbf{A}^*_{nn} \end{bmatrix} = \begin{bmatrix} \mathbf{Det}\,[\mathbf{A}_{ij}] & 0 & \cdots & 0 \\ 0 & \mathbf{Det}\,[\mathbf{A}_{ij}] & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \mathbf{Det}\,[\mathbf{A}_{ij}] \end{bmatrix}.$$

Ist $\mathbf{A}$, also auch $\mathbf{Det}\,[\mathbf{A}_{ij}]$ nichtsingulär, so folgt hieraus unmittelbar

$$(1) \qquad \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ \cdot & \cdots & \cdot \\ \mathbf{A}_{n1} & \cdots & \mathbf{A}_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{A}^*_{11}\,\mathbf{Det}^{-1}\,[\mathbf{A}_{ij}] & \cdots & \mathbf{A}^*_{1n}\,\mathbf{Det}^{-1}\,[\mathbf{A}_{ij}] \\ \cdot & \cdots & \cdot \\ \mathbf{A}^*_{n1}\,\mathbf{Det}^{-1}\,[\mathbf{A}_{ij}] & \cdots & \mathbf{A}^*_{nn}\,\mathbf{Det}^{-1}\,[\mathbf{A}_{ij}] \end{bmatrix} = \mathbf{E}_{mn}.$$

Damit haben wir die Inversion der Matrix $\mathbf{A} = [\mathbf{A}_{ij}]$ nm-ter Ordnung auf Additionen und Multiplikationen und auf die Inversion der einzigen Matrix $\mathbf{Det}\,[\mathbf{A}_{ij}]$ m-ter Ordnung zurückgeführt.

Hierzu ist allerdings zu bemerken, daß die Berechnung der Determinantenmatrix und der Komplementmatrizen im allgemeinen noch langwieriger ist, als die Berechnung einer gewöhnlichen Determinante. Die durch die Formel (1) gegebene Inversionsmethode wird also nur dann eine praktische Brauchbarkeit haben, wenn die darin vorkommenden Matrizen besonders einfach berechenbar sind. Dies ist jedoch bei der Inversion der Laplace—Poissonschen Operatormatrix der Fall.

Wie wir im Abschnitt I gezeigt haben, hat die Laplace—Poissonsche Operatormatrix (für ein Rechteck und verschwindende Randwerte) die folgende Form:

$$\mathbf{L} = \begin{array}{c} \begin{bmatrix} \mathbf{K} & -\mathbf{E} & \mathbf{O} & \cdots & \mathbf{O} \\ -\mathbf{E} & \mathbf{K} & -\mathbf{E} & \cdots & \mathbf{O} \\ \mathbf{O} & -\mathbf{E} & \mathbf{K} & \cdots & \mathbf{O} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \mathbf{O} & \cdot & \cdot & \cdots\ -\mathbf{E} & \mathbf{K} \end{bmatrix} \end{array} \begin{array}{l} (1 \\ (2 \\ (3\,, \\ \\ (n \end{array}$$

$$\mathbf{K} = \begin{bmatrix} 4 & -1 & 0 & \cdots & 0 \\ -1 & 4 & -1 & \cdots & 0 \\ 0 & -1 & 4 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdot & \cdot & \cdots\ -1 & 4 \end{bmatrix} \begin{array}{l} (1 \\ (2 \\ (3\,, \\ \\ (m \end{array} \qquad \mathbf{E} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}.$$

Die entsprechende, aus Skalarelementen bestehende Matrix

$$\mathbf{C}_n(x) = \begin{array}{c} \begin{bmatrix} x & -1 & 0 & \cdots & 0 \\ -1 & x & -1 & \cdots & 0 \\ 0 & -1 & x & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdot & \cdot & \cdots -1 & x \end{bmatrix} \end{array} \begin{array}{l} (1 \\ (2 \\ (3 \\ \\ (n \end{array}$$

ist nichts anderes als die wohlbekannte Kontinuantmatrix, deren Determinante und Inverse explizit bekannt sind. Wird nämlich das Tschebyscheffsche Polynom k-ten Grades (zweiter Art) mit $T_k(x)$ bezeichnet, so besteht zwischen diesen Polynomen die rekurrente Relation

(2) $$T_{k+1}(x) = x\,T_k(x) - T_{k-1}(x); \quad T_1 = x, \quad T_0 = 1,$$

und daraus folgt, daß

$$\operatorname{Det} \mathbf{C}_n(x) = T_n(x)$$

und die Inverse von $\mathbf{C}_n(x)$ in der folgenden expliziten Form angegeben werden kann:

(3) $$\mathbf{C}_n(x)^{-1} = \begin{bmatrix} \frac{T_{n-1}(x)\,T_0(x)}{T_n(x)} & \frac{T_{n-2}(x)\,T_0(x)}{T_n(x)} & \cdots & \frac{T_0^2(x)}{T_n(x)} \\ \frac{T_{n-2}(x)\,T_0(x)}{T_n(x)} & \frac{T_{n-2}(x)\,T_1(x)}{T_n(x)} & \cdots & \frac{T_0(x)\,T_1(x)}{T_n(x)} \\ \cdot & \cdot & \cdots & \cdot \\ \frac{T_0^2(x)}{T_n(x)} & \frac{T_0(x)\,T_1(x)}{T_n(x)} & \cdots & \frac{T_0(x)\,T_{n-1}(x)}{T_n(x)} \end{bmatrix}.$$

Für $x = 2$ ergibt sich z. B. hieraus

$$\mathbf{C}_n(2)^{-1} = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & & \\ & & & \ddots & \\ & & & -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{n\cdot 1}{n+1} & \frac{(n-1)\cdot 1}{n+1} & \cdots & \frac{1\cdot 1}{n+1} \\ \frac{(n-1)\cdot 1}{n+1} & \frac{(n-1)\cdot 2}{n+1} & \cdots & \frac{1\cdot 2}{n+1} \\ \cdot & \cdot & \cdots & \cdot \\ \frac{1\cdot 1}{n+1} & \frac{1\cdot 2}{n+1} & \cdots & \frac{1\cdot n}{n+1} \end{bmatrix}.$$

Werden nun in der Gleichung (3) an die Stelle von $x, 1, 0$ die vertauschbaren Blöcke von $\mathbf{L}$

$$\mathbf{K}, \mathbf{E}, \mathbf{O}$$

substituiert, so erhält man für die Inverse der Laplace—Poissonschen Operatormatrix unmittelbar den folgenden expliziten Ausdruck:

$$
\mathbf{L}^{-1} = \begin{bmatrix} T_{n-1}(\mathbf{K})\,T_0(\mathbf{K})\,T_n(\mathbf{K})^{-1} & T_{n-2}(\mathbf{K})\,T_0(\mathbf{K})\,T_n(\mathbf{K})^{-1} & \cdots & T_0(\mathbf{K})^2\,T_n(\mathbf{K})^{-1} \\ T_{n-2}(\mathbf{K})\,T_0(\mathbf{K})\,T_n(\mathbf{K})^{-1} & T_{n-2}(\mathbf{K})\,T_1(\mathbf{K})\,T_n(\mathbf{K})^{-1} & \cdots & T_0(\mathbf{K})\,T_1(\mathbf{K})\,T_n(\mathbf{K})^{-1} \\ \cdot & \cdot & \cdots & \cdot \\ T_0(\mathbf{K})^2\,T_n(\mathbf{K})^{-1} & T_0(\mathbf{K})\,T_1(\mathbf{K})\,T_n(\mathbf{K})^{-1} & \cdots & T_0(\mathbf{K})\,T_{n-1}(\mathbf{K})\,T_n(\mathbf{K})^{-1} \end{bmatrix}. \tag{4}
$$

Die Matrix **L**, sowie ihre Inverse sind zentrosymmetrisch, deshalb gibt es unter den n^2 Blöcken nur ungefähr $\frac{n^2}{4}$ verschiedene.

Die durch die rekurrenten Relationen (2) definierten Polynome $T_k(x)$ können auch in expliziter Form angegeben werden:

$$
T_2(x) = x^2 - 1, \quad T_3(x) = x^3 - 2x, \quad T_4(x) = x^4 - 3x^2 + 1,
$$

$$
T_k(x) = x^k - \binom{k-1}{1} x^{k-2} + \binom{k-2}{2} x^{k-4} - + \cdots .
$$

Für numerische Rechnungen scheint es aber am zweckmäßigsten, wenn man zuerst mit Hilfe der rekurrenten Relationen (2) die Folge der Matrizen m-ter Ordnung

$$
T_1(\mathbf{K}),\ T_2(\mathbf{K}), \ldots,\ T_{n-1}(\mathbf{K}),\ T_n(\mathbf{K})^{-1},
$$

und dann gemäß (4) die einzelnen Blöcke von $\mathbf{L}^{-1}$ berechnet.

Aus der mechanischen Deutung der Laplace—Poissonschen Koeffizientenmatrix, welche später in dieser Arbeit noch näher auseinandergesetzt wird, folgt unmittelbar, daß alle Elemente der Inversen $\mathbf{L}^{-1}$ positiv sind. Rein mathematisch kann diese Eigenschaft der Inversen bewiesen werden, wenn man eine vom Verfasser herrührende Verallgemeinerung eines Stieltjesschen Satzes heranzieht [1].

Dieser Satz besagt nämlich, daß, wenn

1. alle Hauptminoren einer Matrix positiv sind,
2. alle Elemente außerhalb der Hauptdiagonalen nichtpositiv sind,
3. alle Teilzeilen rechts von der Hauptdiagonalen und alle Teilspalten unterhalb der Hauptdiagonalen mindestens ein negatives Element enthalten,

alle Elemente der Inversen positiv sind.

Diese drei Bedingungen sind jedoch bei der Laplace—Poissonschen Koeffizientenmatrix **L** offenbar erfüllt, folglich hat ihre Inverse lauter positive Elemente.

Zweite Methode

Aus zwei quadratischen Matrizen **A** und **B** beliebiger Ordnung kann man durch Additionen, Multiplikationen, skalare Multiplikation und direkte Multiplikation Polynome folgender Art bilden:

$$\varphi(\mathbf{A}, \mathbf{B}) = \sum_p \sum_q c_{pq} \mathbf{A}^p \cdot \times \mathbf{B}^q .$$

Diese Polynome, die wir kurz direkte Polynome nennen wollen, wurden zuerst von C. STÉPHANOS untersucht [4]. Von ihm stammt der folgende Satz über die Eigenwerte von $\varphi(\mathbf{A}, \mathbf{B})$:

Sind die Eigenwerte der Matrix **A** n-ter Ordnung

$$a_1, a_2, \ldots, a_n$$

und die Eigenwerte der Matrix **B** m-ter Ordnung

$$b_1, b_2, \ldots, b_m,$$

dann sind die Eigenwerte des direkten Polynoms $\varphi(\mathbf{A}, \mathbf{B})$

$$\varphi(a_1, b_1), \ldots, \varphi(a_1, b_m), \varphi(a_2, b_1), \ldots, \varphi(a_2, b_m), \ldots, \varphi(a_n, b_1), \ldots, \varphi(a_n, b_m). \tag{5}$$

Als eine naheliegende Ergänzung der Ergebnisse von STÉPHANOS verifiziert man unmittelbar, daß im Besitze der Spektralzerlegung von $\mathbf{C}_n$ und $\mathbf{C}_m$ man auch die Spektralzerlegung für $\varphi(\mathbf{C}_n, \mathbf{C}_m)$, folglich auch für $\varphi(\mathbf{C}_n, \mathbf{C}_m)^{-1}$ aufschreiben kann.

Die Spektralzerlegung für $\mathbf{C}_n$ ist bekanntlich

$$\mathbf{C}_n = \mathbf{U}_n \mathbf{\Lambda}_n \mathbf{U}_n, \quad \mathbf{U}_n = \mathbf{U}_n', \quad \mathbf{U}_n \mathbf{U}_n' = \mathbf{E},$$

wo $\mathbf{U}_n$ die aus den Eigenvektoren von $\mathbf{C}_n$ gebildete Orthogonalmatrix

$$\mathbf{U}_n = \sqrt{\frac{2}{n+1}} \begin{bmatrix} \sin\frac{\pi}{n+1} & \sin\frac{2\pi}{n+1} & \cdots & \sin\frac{n\pi}{n+1} \\ \sin\frac{2\pi}{n+1} & \sin\frac{4\pi}{n+1} & \cdots & \sin\frac{2n\pi}{n+1} \\ \cdot & \cdot & \cdots & \cdot \\ \sin\frac{n\pi}{n+1} & \sin\frac{2n\pi}{n+1} & \cdots & \sin\frac{n^2\pi}{n+1} \end{bmatrix} \tag{6}$$

und $\mathbf{\Lambda}_n$ die aus den Eigenwerten

$$\lambda_p^{(n)} = 4 \sin^2 \frac{p\pi}{2(n+1)} \tag{7}$$

von $\mathbf{C}_n$ gebildete Diagonalmatrix bedeutet.

Hieraus sieht man leicht ein, daß die Spektralzerlegung von $\varphi(\mathbf{C}_n, \mathbf{C}_m)$ folgende Form besitzt:

$$\varphi(\mathbf{C}_n, \mathbf{C}_m) = \sum_p \sum_q c_{pq} \mathbf{C}_n^p \cdot \times \mathbf{C}_m^q = \mathbf{U}_n \cdot \times \mathbf{U}_m \langle \varphi(\lambda_p^{(n)}, \lambda_q^{(m)}) \rangle \mathbf{U}_n \cdot \times \mathbf{U}_m,$$

wo $\langle \varphi(\lambda_p^{(n)} \lambda_q^{(m)}) \rangle$ die aus den nm Eigenwerten (7) gebildete Diagonalmatrix bedeutet.

Also erhalten wir für die Matrix $\mathbf{L} = \mathbf{C}_n \cdot \times \mathbf{E}_m + \mathbf{E}_n \cdot \times \mathbf{C}_m$

$$\mathbf{L} = \mathbf{U}_n \cdot \times \mathbf{U}_m \langle \lambda_p^{(n)} + \lambda_q^{(m)} \rangle \mathbf{U}_n \cdot \times \mathbf{U}_m,$$

und für die Inverse $\mathbf{L}^{-1}$ ergibt sich hieraus

$$\mathbf{L}^{-1} = \mathbf{U}_n \cdot \times \mathbf{U}_m \left\langle \frac{1}{\lambda_p^{(n)} + \lambda_q^{(m)}} \right\rangle \mathbf{U}_n \cdot \times \mathbf{U}_m .$$

Wie die Formeln (6) und (7) zeigen, können sämtliche in diesen Formeln vorkommenden Zahlen aus goniometrischen Tafeln unmittelbar entnommen werden.

III. Berücksichtigung einer einfachen Modifikation der Randbedingungen

Um den Einfluß einer Modifikation der Randbedingungen auf die Lösung der Poissonschen Differenzengleichung anschaulich darstellen zu können, geben wir zuerst der Aufgabe eine mechanische Interpretation.

Werden in den Knotenpunkten P_{ij} eines quadratischen elastischen Fadennetzes die transversalen Kräfte q_{ij} angebracht, während die Endpunkte

$$P_{i0},\ P_{i,m+1},\ P_{0j},\ P_{n+1,j} \qquad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, m)$$

der Fäden festgehalten werden, so werden die transversalen Verschiebungen x_{ij}, der Knotenpunkte P_{ij} (bei geeigneter Wahl der Einheiten) durch die Gleichgewichtsgleichungen

$$4x_{ij} - x_{i+1,j} - x_{i-1,j} - x_{i,j-1} - x_{i,j+1} = q_{ij} \tag{8}$$
$$(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, m)$$

bestimmt, wo gemäß den Randbedingungen

$$x_{0j},\ x_{n+1,j},\ x_{i0},\ x_{i,m+1}$$

veschwinden müssen.

Es sei jetzt der Knotenpunkt P_{rs} festgehalten. Dies bedeutet erstens daß die Verschiebung $x_{rs} = 0$ gesetzt werden muß. Zweitens bekommt die

Gleichung mit dem Doppelindex (r, s) eine neue Bedeutung. Die Festhaltung des Knotenpunktes P_{rs} kann man nämlich auch in der Weise realisiert denken, daß man im Knotenpunkte P_{rs} eine vorläufig unbekannte, von den anderen Kräften q_{ij} abhängige Kraft q_{rs} anbringt, welche die Verschiebung x_{rs} ausgleicht. Diese Gleichung kann also nachträglich zur Berechnung der beim Festhalten des Knotenpunktes P_{rs} auftretenden Reaktionskraft benutzt werden, nachdem man erst die unbekannten Verschiebungen x_{ij} $((i, j) \neq (r, s))$ aus den übrigbleibenden $nm-1$ Gleichungen (8) berechnet hat.

Die Koeffizientenmatrix dieser Gleichungen ist aber nichts anderes als diejenige (gleichfalls symmetrische und definite) Hauptminormatrix, welche aus $\mathbf{L}$ durch Tilgung der Zeile und der Spalte mit dem Doppelindex (r, s) entsteht.

Wir werden auf diese Weise zur Aufgabe geführt, einen einfachen Algorithmus anzugeben, welche die Berechnung der Inversen einer Minormatrix aus der schon bekannten Inversen der ursprünglichen Matrix ermöglicht. Das geeignete Hilfsmittel zur Lösung dieser Aufgabe ist, wie wir jetzt zeigen wollen, eine besonders einfache rangvermindernde Operation.

Es sei $\mathbf{A}$ eine beliebige Matrix. Wir betrachten die folgende, aus $\mathbf{A}$ abgeleitete Matrix:

$$\mathbf{A}_1 = \mathbf{A} - \frac{\mathbf{A u v' A}}{\mathbf{v' A u}}, \tag{9}$$

wo $\mathbf{u}$ und $\mathbf{v}'$ bis auf die Einschränkung $\mathbf{v'Au} \neq 0$ beliebige Parametervektoren sind. Es ist leicht zu zeigen, daß sich der Rang von $\mathbf{A}$ bei dieser Operation genau um Eins vermindert, d. h.

$$\operatorname{Rang}\left(\mathbf{A} - \frac{\mathbf{A u v' A}}{\mathbf{v' A u}}\right) = \operatorname{Rang}(\mathbf{A}) - 1.$$

Es sei, wie üblich, der i-te Einheitsspaltenvektor mit $\mathbf{e}_i$, und der j-te Einheitszeilenvektor mit $\mathbf{e}_j'$ bezeichnet. Ist das Element $a_{ij} = \mathbf{e}_j' \mathbf{A} \mathbf{e}_i$ von $\mathbf{A}$ von 0 verschieden, so kann $\mathbf{u} = \mathbf{e}_i$, $\mathbf{v}' = \mathbf{e}_j'$ gesetzt werden. Bei dieser Wahl der Parametervektoren ist jedoch $\mathbf{A e}_i$ die i-te Spalte von $\mathbf{A}$, und $\mathbf{e}_j' \mathbf{A}$ ist die j-te Zeile von $\mathbf{A}$. Die Matrix vom Typ (9), welche bei dieser Wahl der Parametervektoren erhalten wird, erfüllt also offenbar die Relationen

$$\left(\mathbf{A} - \frac{\mathbf{A e}_i \mathbf{e}_j' \mathbf{A}}{\mathbf{e}_j' \mathbf{A e}_i}\right) \mathbf{e}_i = 0, \quad \mathbf{e}_j' \left(\mathbf{A} - \frac{\mathbf{A e}_i \mathbf{e}_j' \mathbf{A}}{\mathbf{e}_j' \mathbf{A e}_i}\right) = 0,$$

d. h. die i-te Spalte und die j-te Zeile der abgeleiteten Matrix bestehen aus lauter Nullen.

Nach dieser Vorbereitung beweisen wir den folgenden

SATZ. *Wenn man an der Inversen* $\mathbf{A}^{-1}$ *einer Matrix* $\mathbf{A}$ *N-ter Ordnung die rangvermindernde Operation mit Hilfe der Formel*

$$\mathbf{A}^{-1}-\frac{\mathbf{A}^{-1}\mathbf{e}_i\mathbf{e}_j'\mathbf{A}^{-1}}{\mathbf{e}_j'\mathbf{A}^{-1}\mathbf{e}_i} \tag{10}$$

ausführt, und die so entstehende i-te Spalte und j-te Zeile (welche aus lauter 0*-en bestehen) wegläßt, die übrigbleibende Matrix N*—1*-ter Ordnung gleich der Inversen derjenigen Minormatrix von* $\mathbf{A}$ *ist, welche durch Tilgung der j-ten Spalte und der i-ten Zeile entsteht.*[2]

BEWEIS. Werden die Elemente der Inversen durch ihre wohlbekannten Ausdrücke A_{ij}/A ersetzt, wo $A=\operatorname{Det}\mathbf{A}$ und A_{ij} das algebraische Komplement von a_{ji} bedeutet, so erhalten die Elemente der Matrix (10) die folgende Form:

$$\frac{A_{pq}}{A}-\frac{A_{pi}A_{jq}}{AA_{ji}}=\frac{A_{ji}A_{pq}-A_{pi}A_{jq}}{AA_{ji}}.$$

Die im Zähler vorkommende Determinante zweiter Ordnung ist aber nach einem Satz von JACOBI ([5], S. 61) gleich dem Produkt $AA_{ji,pq}$, wo $A_{ji,pq}$ die zu den Indexpaaren ij, qp gehörige Unterdeterminante $N-2$-ter Ordnung von $\mathbf{A}$ bedeutet. Das allgemeine Element von (10) ist also

$$\frac{A_{ji,pq}}{A_{ji}},$$

d. h. genau das zum Indexpaar pq gehörige Element von $\mathbf{A}_{ij}^{-1}$. Q. e. d.

Aus diesem Ergebnis folgt, daß die Hinzunahme einer neuen Bedingung $x_{r\varrho}=0$ zu den schon vorhandenen Randbedingungen dadurch berücksichtigt werden kann, daß man die Inverse der ursprünglichen Matrix einer rangvermindernden Operation unterwirft. Werden mehrere, z. B. k neue derartige Randbedingungen vorgeschrieben, so kann man diese neuen Randbedingungen durch die sukzessive Ausführung von k Operationen von der Form (10) berücksichtigen.

Im folgenden Abschnitt soll noch gezeigt werden, daß die k neuen Rand-

[2] *Bemerkung des Redakteurs.* Dasselbe Resultat haben, unabhängig von E. EGERVÁRY, P. RÓZSA und N. SIEBER erhalten. (S. П. Рожа, О применении клеточных матриц в механике корпускулярных систем, Усп. Мат. Наук., 14 (1959), вып. 4 (88), S. 207—211; H. STENKER und N. SIEBER, Ein Reduktionssatz über Umkehrmatrizen und seine Anwendung auf ein Beispiel aus der Statik, *Wissenschaftliche Zeitschrift der Hochschule für Architektur und Bauwesen Weimar,* 6 (1958/59), H. 2, S. 105—117.)

bedingungen auch durch eine einzige, „mehrfache" rangvermindernde Operation berücksichtigt werden können, wobei allerdings die Inversion einer Matrix k-ter Ordnung ausgeführt werden muß.

IV. Berücksichtigung einer mehrfachen Modifikation der Randbedingungen

In diesem Abschnitt werden wir zeigen, daß das Ergebnis $\mathbf{A}_\varkappa$ von $\varkappa$ sukzessiv ausgeführten einfachen rangvermindernden Operationen (9) auch durch die folgende „mehrfache" rangvermindernde Operation gewonnen werden kann:

$$\mathbf{A}_\varkappa = \mathbf{A} - \mathbf{A}\mathbf{U}_\varkappa(\mathbf{V}'_\varkappa \mathbf{A}\mathbf{U}_\varkappa)^{-1}\mathbf{V}'_\varkappa \mathbf{A}. \tag{11}$$

Hier ist

$$\mathbf{U}_\varkappa = [\mathbf{u}_1, \ldots, \mathbf{u}_\varkappa], \quad \mathbf{V}'_\varkappa = \begin{bmatrix} \mathbf{v}'_1 \\ \mathbf{v}'_2 \\ \vdots \\ \mathbf{v}'_\varkappa \end{bmatrix}, \quad |\mathbf{V}'_\varkappa \mathbf{A}\mathbf{U}_\varkappa| \neq 0.$$

Nehmen wir an, daß (11) schon für $\varkappa = 1, 2, \ldots, k$ bewiesen ist. Dann erhalten wir durch Ausführung einer $k+1$-ten rangvermindernden Operation

$$\mathbf{A}_{k+1} = \mathbf{A}_k - \frac{\mathbf{A}_k \mathbf{u}_{k+1} \mathbf{v}'_{k+1} \mathbf{A}_k}{\mathbf{v}'_{k+1} \mathbf{A}_k \mathbf{u}_{k+1}}.$$

Wird jetzt hier der Ausdruck (11) von $\mathbf{A}_k$ substituiert, so folgt

$$\mathbf{A}_{k+1} = \mathbf{A} - \mathbf{A}\mathbf{U}_k(\mathbf{V}'_k\mathbf{A}\mathbf{U}_k)^{-1}\mathbf{V}'_k\mathbf{A} -$$
$$- \frac{\{\mathbf{A} - \mathbf{A}\mathbf{U}_k(\mathbf{V}'_k\mathbf{A}\mathbf{U}_k)^{-1}\mathbf{V}'_k\mathbf{A}\}\mathbf{u}_{k+1}\mathbf{v}'_{k+1}\{\mathbf{A} - \mathbf{A}\mathbf{U}_k(\mathbf{V}'_k\mathbf{A}\mathbf{U}_k)^{-1}\mathbf{V}'_k\mathbf{A}\}}{\mathbf{v}'_{k+1}\{\mathbf{A} - \mathbf{A}\mathbf{U}_k(\mathbf{V}'_k\mathbf{A}\mathbf{U}_k)^{-1}\mathbf{V}'_k\mathbf{A}\}\mathbf{u}_{k+1}}.$$

Wir werden zeigen, daß die Formel (11) für $\varkappa = k+1$ zu demselben Ergebnis führt. Wir schreiben zuerst $\mathbf{U}_{k+1}$ und $\mathbf{V}'_{k+1}$ als partitionierte Matrizen:

$$\mathbf{U}_{k+1} = [\mathbf{U}_k \mathbf{u}_{k+1}], \quad \mathbf{V}'_{k+1} = \begin{bmatrix} \mathbf{V}'_k \\ \mathbf{v}'_{k+1} \end{bmatrix}, \quad \mathbf{V}'_{k+1}\mathbf{A}\mathbf{U}_{k+1} = \begin{bmatrix} \mathbf{V}'_k\mathbf{A}\mathbf{U}_k & \mathbf{V}'_k\mathbf{A}\mathbf{u}_{k+1} \\ \mathbf{v}'_{k+1}\mathbf{A}\mathbf{U}_k & \mathbf{v}'_{k+1}\mathbf{A}\mathbf{u}_{k+1} \end{bmatrix}$$

Werden diese Ausdrücke in (11) substituiert, so erhält die Formel (11) für $\varkappa = k+1$ folgende Form:

$$\mathbf{A}_{k+1} = \mathbf{A} - \mathbf{A}[\mathbf{U}_k \mathbf{u}_{k+1}] \begin{bmatrix} \mathbf{V}'_k\mathbf{A}\mathbf{U}_k & \mathbf{V}'_k\mathbf{A}\mathbf{u}_{k+1} \\ \mathbf{v}'_{k+1}\mathbf{A}\mathbf{U}_k & \mathbf{v}'_{k+1}\mathbf{A}\mathbf{u}_{k+1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{V}'_k \\ \mathbf{v}'_{k+1} \end{bmatrix} \mathbf{A}. \tag{12}$$

Durch Anwendung einer bekannten Formel für die Inversion einer partitionier-

ten Matrix bekommt man hieraus

$$\begin{bmatrix} \mathbf{V}_k'\mathbf{A}\mathbf{U}_k & \mathbf{V}_k'\mathbf{A}\mathbf{u}_{k+1} \\ \mathbf{v}_{k+1}'\mathbf{A}\mathbf{U}_k & \mathbf{v}_{k+1}'\mathbf{A}\mathbf{u}_{k+1} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{V}_k'\mathbf{A}\mathbf{U}_k)^{-1} & 0 \\ 0 & 0 \end{bmatrix} +$$

$$+\frac{1}{\mathbf{v}_{k+1}'\{\mathbf{A}-\mathbf{A}\mathbf{U}_k(\mathbf{V}_k'\mathbf{A}\mathbf{U}_k)^{-1}\mathbf{V}_k'\mathbf{A}\}\mathbf{u}_{k+1}} \begin{bmatrix} (\mathbf{V}_k'\mathbf{A}\mathbf{U}_k)^{-1}\mathbf{V}_k'\mathbf{A}\mathbf{u}_{k+1} \\ -1 \end{bmatrix} [\mathbf{v}_{k+1}'\mathbf{A}\mathbf{U}_k(\mathbf{V}_k'\mathbf{A}\mathbf{U}_k)^{-1} \quad -1].$$

Wird dieses Ergebnis in (12) eingesetzt, so erkennt man unmittelbar, daß beide Ausdrücke für $\mathbf{A}_{k+1}$ übereinstimmen. Damit haben wir die Gültigkeit von (11) durch Induktion bewiesen.

Wir schreiben jetzt die Inverse der symmetrischen, positiv-definiten Matrix $\mathbf{L}$ in der Form

$$\mathbf{L}^{-1} = \mathbf{R} = [\mathbf{r}_1 \dots \mathbf{r}_N] = \begin{bmatrix} \mathbf{r}_1' \\ \vdots \\ \mathbf{r}_N' \end{bmatrix} = [r_{ij}] \qquad (i, j = 1, 2, \dots, N),$$

und stellen uns die Aufgabe, die Inverse der Hauptminormatrix $N-k$-ter Ordnung, welche aus $\mathbf{L}$ durch Tilgung der i_1-ten, i_2-ten, ..., i_k-ten Zeilen und Spalten entsteht, mit Hilfe von $\mathbf{L}^{-1}$ zu berechnen. Zu diesem Zwecke setzen wir in der allgemeinen Formel (11)

$$\mathbf{u}_1 = \mathbf{e}_{i_1}, \ \mathbf{u}_2 = \mathbf{e}_{i_2}, \dots, \ \mathbf{u}_k = \mathbf{e}_{i_k},$$
$$\mathbf{v}_1' = \mathbf{e}_{i_1}', \ \mathbf{v}_2' = \mathbf{e}_{i_2}', \dots, \ \mathbf{v}_k' = \mathbf{e}_{i_k}'.$$

Dann wird

$$\mathbf{R}\mathbf{U}_k = [\mathbf{r}_{i_1}\mathbf{r}_{i_2} \dots \mathbf{r}_{i_k}], \quad \mathbf{V}_k'\mathbf{R} = \begin{bmatrix} \mathbf{r}_{i_1}' \\ \mathbf{r}_{i_2}' \\ \vdots \\ \mathbf{r}_{i_k}' \end{bmatrix},$$

und

$$\mathbf{V}_k'\mathbf{R}\mathbf{U}_k = \begin{bmatrix} r_{i_1 i_1} & \dots & r_{i_1 i_k} \\ \cdot & \cdots & \cdot \\ r_{i_k i_1} & \dots & r_{i_k i_k} \end{bmatrix}$$

ist invertierbar.

Wird jetzt die k-fache rangvermindernde Operation

$$\mathbf{R} - [\mathbf{r}_{i_1} \dots \mathbf{r}_{i_k}] \begin{bmatrix} r_{i_1 i_1} & \dots & r_{i_1 i_k} \\ \cdot & \cdots & \cdot \\ r_{i_k i_1} & \dots & r_{i_k i_k} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{r}_{i_1}' \\ \vdots \\ \mathbf{r}_{i_k}' \end{bmatrix}$$

ausgeführt, so ergibt sich eine Matrix, deren i_1-, i_2-, ..., i_k-te Zeilen und Spalten aus lauter 0-en bestehen. Diejenige Matrix $N-k$-ter Ordnung, welche

nach Tilgung dieser 0-Zeilen und 0-Spalten übrigbleibt, ist nichts anderes als die gesuchte Inverse der oben ausgewählten Hauptminormatrix.

Auf diese Weise haben wir ein Verfahren gefunden, welche die Berücksichtigung von k neuen Randbedingungen durch eine einzige und explizite matrizentechnische Operation gestattet.

V. Beispiele

Wir betrachten jetzt ein quadratisches Netz mit den 9 inneren Knotenpunkten und mit den Randpunkten

$$P_{ij} \qquad (i, j = 0, 1, 2, 3, 4).$$

Bei verschwindenden Randwerten und bei lexikographischer Anordnung hat das Poissonsche Differenzengleichungssystem folgende Form:

$$\begin{array}{llllll} 4x_{11} - x_{12} & & -x_{21} & & & = p_{11}, \\ -x_{11} + 4x_{12} - x_{13} & & & -x_{22} & & = p_{12}, \\ & & & & & \vdots \\ & & & -x_{23} & -x_{32} + 4x_{33} & = p_{33}. \end{array}$$

Wird die Koeffizientenmatrix $\mathbf{L}$ 9-ter Ordnung dieses Gleichungssystems gemäß den Ausführungen des Abschnittes I folgendermaßen partitioniert:

$$\mathbf{L} = \begin{bmatrix} \mathbf{K} & -\mathbf{E} & \mathbf{O} \\ -\mathbf{E} & \mathbf{K} & -\mathbf{E} \\ \mathbf{O} & -\mathbf{E} & \mathbf{K} \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix},$$

so erhält man aus der Formel (4)

$$\mathbf{L}^{-1} = \begin{bmatrix} T_2(\mathbf{K})\,T_0(\mathbf{K})\,T_3(\mathbf{K})^{-1} & T_1(\mathbf{K})\,T_0(\mathbf{K})\,T_3(\mathbf{K})^{-1} & T_0(\mathbf{K})^2\,T_3(\mathbf{K})^{-1} \\ T_1(\mathbf{K})\,T_0(\mathbf{K})\,T_3(\mathbf{K})^{-1} & T_1(\mathbf{K})^2\,T_3(\mathbf{K})^{-1} & T_0(\mathbf{K})\,T_1(\mathbf{K})\,T_3(\mathbf{K})^{-1} \\ T_0(\mathbf{K})^2\,T_3(\mathbf{K})^{-1} & T_0(\mathbf{K})\,T_1(\mathbf{K})\,T_3(\mathbf{K})^{-1} & T_0(\mathbf{K})\,T_2(\mathbf{K})\,T_3(\mathbf{K})^{-1} \end{bmatrix}. \tag{13}$$

Aus der Rekursionsformel (2) folgt

$$T_0(\mathbf{K}) = \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T_1(\mathbf{K}) = \mathbf{K} = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix},$$

$$T_2(\mathbf{K}) = \mathbf{K}^2 - \mathbf{E} = \begin{bmatrix} 16 & -8 & 1 \\ -8 & 17 & -8 \\ 1 & -8 & 16 \end{bmatrix}, \quad T_3(\mathbf{K}) = \mathbf{K}^3 - 2\mathbf{K} = 4\begin{bmatrix} 17 & -12 & 3 \\ -12 & 20 & -12 \\ 3 & -12 & 17 \end{bmatrix},$$

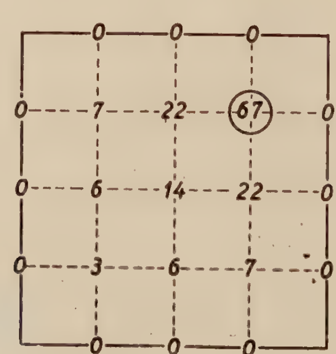
$$T_3(\mathbf{K})^{-1} = \frac{1}{224}\begin{bmatrix} 7 & 6 & 3 \\ 6 & 10 & 6 \\ 3 & 6 & 7 \end{bmatrix}.$$

Substituiert man diese Matrizen in (13), so ergibt sich

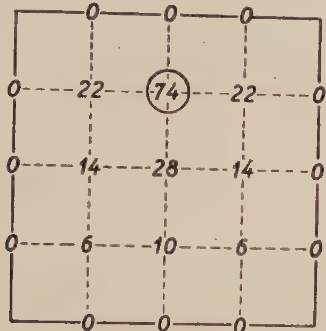
$$\mathbf{R}=\mathbf{L}^{-1}=\left[\begin{array}{rrr|rrr|rrr} 16 & -8 & 1 & 4 & -1 & 0 & 1 & 0 & 0 \\ -8 & 17 & -8 & -1 & 4 & -1 & 0 & 1 & 0 \\ 1 & -8 & 16 & 0 & -1 & 4 & 0 & 0 & 1 \\ \hline 4 & -1 & 0 & 17 & -8 & 1 & 4 & -1 & 0 \\ -1 & 4 & -1 & -8 & 18 & -8 & -1 & 4 & -1 \\ 0 & -1 & 4 & 1 & -8 & 17 & 0 & -1 & 4 \\ \hline 1 & 0 & 0 & 4 & -1 & 0 & 16 & -8 & 1 \\ 0 & 1 & 0 & -1 & 4 & -1 & -8 & 17 & -8 \\ 0 & 0 & 1 & 0 & -1 & 4 & 1 & -8 & 16 \end{array}\right]\cdot\times\frac{\begin{bmatrix} 7 & 6 & 3 \\ 6 & 10 & 6 \\ 3 & 6 & 7 \end{bmatrix}}{224}=$$

$$=\frac{1}{224}\left[\begin{array}{rrr|rrr|rrr} 67 & 22 & 7 & 22 & 14 & 6 & 7 & 6 & 3 \\ 22 & 74 & 22 & 14 & 28 & 14 & 6 & 10 & 6 \\ 7 & 22 & 67 & 6 & 14 & 22 & 3 & 6 & 7 \\ \hline 22 & 14 & 6 & 74 & 28 & 10 & 22 & 14 & 6 \\ 14 & 28 & 14 & 28 & 84 & 28 & 14 & 28 & 14 \\ 6 & 14 & 22 & 10 & 28 & 74 & 6 & 14 & 22 \\ \hline 7 & 6 & 3 & 22 & 14 & 6 & 67 & 22 & 7 \\ 6 & 10 & 6 & 14 & 28 & 14 & 22 & 74 & 22 \\ 3 & 6 & 7 & 6 & 14 & 22 & 7 & 22 & 67 \end{array}\right].$$

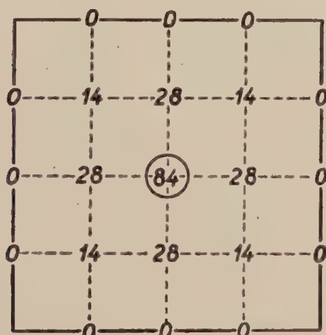
Dieses Ergebnis läßt sich am besten veranschaulichen, wenn man von der lexikographischen Anordnung der Knotenpunkte zur natürlichen quadratischen Anordnung zurückkehrt. So erkennt man, daß die Wertsysteme der Unbekannten, welche aus den ersten, dritten, siebten und neunten Zeilen (oder Spalten) von $\mathbf{L}^{-1}$ entstehen, Spiegelbilder von einander sind, und zwar erhält man aus der dritten Zeile (nach Weglassen des gemeinsamen Faktors 1/224):



Die zweite Zeile ergibt das Wertsystem



(aus der vierten, sechsten, achten Zeile erhält man die Spiegelbilder dieser Verteilung). Die fünfte Zeile (welche mechanisch der Belastung des Fadennetzes in seinem Mittelpunkt entspricht) ergibt das Wertsystem



Jetzt soll die neue Randbedingung $x_{11} = 0$ hinzugenommen werden, das bedeutet, daß der Knotenpunkt P_{11} festgehalten wird. Zum Laplaceschen Operator, welcher den modifizierten Randbedingungen entspricht, gehört die quadratische Minormatrix $\mathbf{L}^*$ 8-ter Ordnung, welche aus $\mathbf{L}$ durch Tilgung der ersten Zeile und ersten Spalte entsteht. Die Inverse dieser Minormatrix läßt sich aus $\mathbf{L}^{-1}$ nach Abschnitt IV mit Hilfe der rangvermindernden Operation

(10) bestimmen:

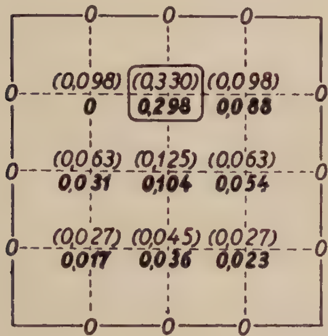
$$\mathbf{L}^{*(-1)} = \mathbf{R} - \frac{\mathbf{r}_{11}\mathbf{r}'_{11}}{r_{11,11}} = \mathbf{R} - \frac{1}{67\cdot 224}\begin{bmatrix}67\\22\\7\\22\\14\\6\\7\\6\\3\end{bmatrix}[67\ \ 22\ \ 7\ \ 22\ \ 14\ \ 6\ \ 7\ \ 6\ \ 3] =$$

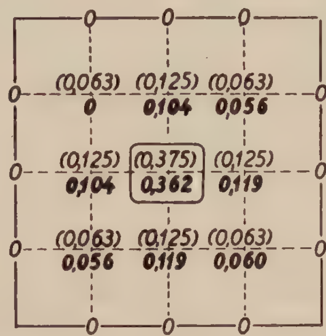
$$= \left[\begin{array}{ccc|ccc|ccc}
0{,}299 & 0{,}098 & 0{,}031 & 0{,}098 & 0{,}063 & 0{,}027 & 0{,}031 & 0{,}027 & 0{,}013 \\
0{,}098 & 0{,}330 & 0{,}098 & 0{,}063 & 0{,}125 & 0{,}063 & 0{,}027 & 0{,}045 & 0{,}027 \\
0{,}031 & 0{,}098 & 0{,}299 & 0{,}027 & 0{,}063 & 0{,}098 & 0{,}013 & 0{,}027 & 0{,}031 \\
\hline
0{,}098 & 0{,}063 & 0{,}027 & 0{,}330 & 0{,}125 & 0{,}045 & 0{,}098 & 0{,}063 & 0{,}027 \\
0{,}063 & 0{,}125 & 0{,}063 & 0{,}125 & 0{,}375 & 0{,}125 & 0{,}063 & 0{,}125 & 0{,}063 \\
0{,}027 & 0{,}063 & 0{,}098 & 0{,}045 & 0{,}125 & 0{,}330 & 0{,}027 & 0{,}063 & 0{,}098 \\
\hline
0{,}031 & 0{,}027 & 0{,}013 & 0{,}098 & 0{,}063 & 0{,}027 & 0{,}299 & 0{,}098 & 0{,}031 \\
0{,}027 & 0{,}045 & 0{,}027 & 0{,}063 & 0{,}125 & 0{,}063 & 0{,}098 & 0{,}330 & 0{,}098 \\
0{,}013 & 0{,}027 & 0{,}031 & 0{,}027 & 0{,}063 & 0{,}098 & 0{,}031 & 0{,}098 & 0{,}299
\end{array}\right] -$$

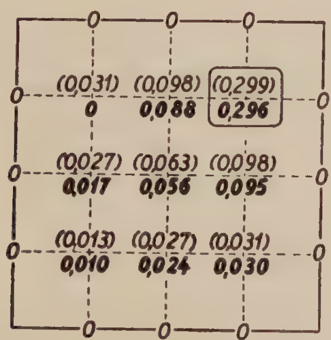
$$- \left[\begin{array}{ccc|ccc|ccc}
0{,}299 & 0{,}098 & 0{,}031 & 0{,}098 & 0{,}063 & 0{,}027 & 0{,}031 & 0{,}027 & 0{,}013 \\
0{,}098 & 0{,}032 & 0{,}010 & 0{,}032 & 0{,}021 & 0{,}009 & 0{,}010 & 0{,}009 & 0{,}004 \\
0{,}031 & 0{,}010 & 0{,}003 & 0{,}010 & 0{,}007 & 0{,}003 & 0{,}003 & 0{,}003 & 0{,}001 \\
\hline
0{,}098 & 0{,}032 & 0{,}010 & 0{,}032 & 0{,}021 & 0{,}009 & 0{,}010 & 0{,}009 & 0{,}004 \\
0{,}063 & 0{,}021 & 0{,}007 & 0{,}021 & 0{,}013 & 0{,}006 & 0{,}007 & 0{,}006 & 0{,}003 \\
0{,}027 & 0{,}009 & 0{,}003 & 0{,}009 & 0{,}006 & 0{,}002 & 0{,}003 & 0{,}002 & 0{,}001 \\
\hline
0{,}031 & 0{,}010 & 0{,}003 & 0{,}010 & 0{,}007 & 0{,}003 & 0{,}003 & 0{,}003 & 0{,}001 \\
0{,}027 & 0{,}009 & 0{,}003 & 0{,}009 & 0{,}006 & 0{,}002 & 0{,}003 & 0{,}002 & 0{,}001 \\
0{,}013 & 0{,}004 & 0{,}001 & 0{,}004 & 0{,}003 & 0{,}001 & 0{,}001 & 0{,}001 & 0{,}001
\end{array}\right] =$$

$$= \left[\begin{array}{ccc|ccc|ccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0{,}298 & 0{,}088 & 0{,}031 & 0{,}104 & 0{,}054 & 0{,}017 & 0{,}036 & 0{,}023 \\
0 & 0{,}088 & 0{,}296 & 0{,}017 & 0{,}056 & 0{,}095 & 0{,}010 & 0{,}024 & 0{,}030 \\
\hline
0 & 0{,}031 & 0{,}017 & 0{,}298 & 0{,}104 & 0{,}036 & 0{,}088 & 0{,}054 & 0{,}023 \\
0 & 0{,}104 & 0{,}056 & 0{,}104 & 0{,}362 & 0{,}119 & 0{,}056 & 0{,}119 & 0{,}060 \\
0 & 0{,}054 & 0{,}095 & 0{,}036 & 0{,}119 & 0{,}328 & 0{,}024 & 0{,}061 & 0{,}097 \\
\hline
0 & 0{,}017 & 0{,}010 & 0{,}088 & 0{,}056 & 0{,}024 & 0{,}296 & 0{,}095 & 0{,}030 \\
0 & 0{,}036 & 0{,}024 & 0{,}054 & 0{,}119 & 0{,}061 & 0{,}095 & 0{,}328 & 0{,}097 \\
0 & 0{,}023 & 0{,}030 & 0{,}023 & 0{,}060 & 0{,}097 & 0{,}030 & 0{,}097 & 0{,}298
\end{array}\right].$$

Wenn man von der lexikographischen Anordnung der Knotenpunkte wieder zur natürlichen quadratischen Anordnung zurückkehrt, so erhält man aus der zweiten, dritten bzw. fünften Zeile die folgenden Wertsysteme (die in Klammern stehenden Werte gehören zu den ursprünglichen Randbedingungen; der Einfluß der modifizierten Randbedingung ist deutlich zu bemerken):







Das vorgeführte Beispiel war geeignet, die einzelnen Operationen zu illustrieren, konnte jedoch wegen der kleinen Ordnungszahlen der darin vorkommenden Matrizen kaum einen Einblick in die praktische Anwendbarkeit der Methode bieten. Deshalb wollen wir hier noch zum Schluß einen Überblick der Rechnungen geben, welche notwendig sind, wenn man ein hinreichend detailliertes Bild des gesuchten funktionalen Zusammenhanges sich verschaffen will.

Nehmen wir an, daß wir die Poissonsche Differenzengleichung für ein L-förmiges Gebiet lösen wollen (Fig. 1), welches 75 innere Knotenpunkte enthält. Die direkte Inversion einer Matrix 75-ter Ordnung ist schon eine Re-

chenaufgabe, welche auch die Fähigkeit vieler Rechenautomaten übertrifft. Unsere Methode erlaubt es, die gesuchte Inverse mit Hilfe einer Rechenmaschine, welche rationale Matrixoperationen bis zur Ordnung 10 ausführen kann (z. B. des Types Pegasus der Firme Ferranti), ohne Schwierigkeiten zu berechnen.

Wir wollen zuerst die Laplacesche Operatormatrix invertieren, welche zum „einfassenden" quadratischen Gebiet mit $10 \times 10 = 100$ inneren Knotenpunkten gehört. Zu diesem Zwecke muß man mit Hilfe der Rekursionsfor-

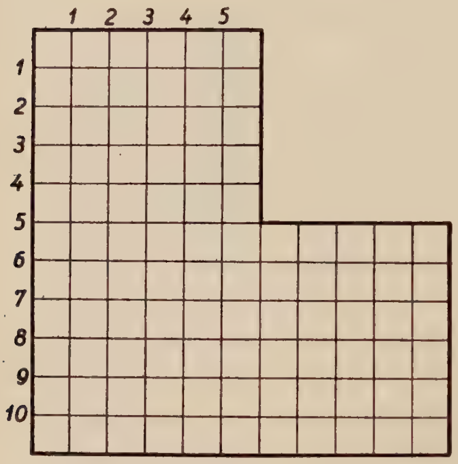


Fig. 1

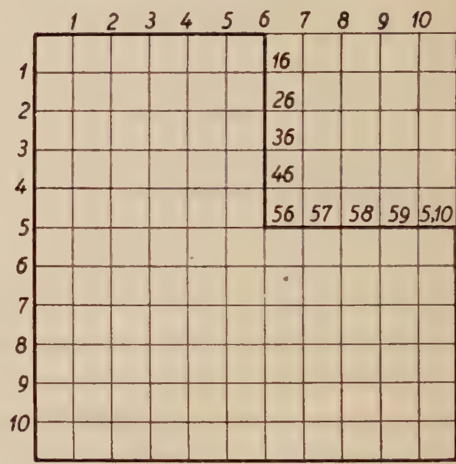


Fig. 2

meln (Abschnitt II, (2)) folgende Matrixpolynome der Matrix $\mathbf{K}$ 10-ter Ordnung berechnen:

$$T_0(\mathbf{K}) = \mathbf{E}, \; T_1(\mathbf{K}) = \mathbf{K}, \ldots, \; T_{11}(\mathbf{K}),$$

und dann die Matrix $T_{11}(\mathbf{K})$ 10-ter Ordnung invertieren. Im Besitze dieser Matrizen erhalten wir die 100 Blöcke von $\mathbf{L}^{-1} = \mathbf{R}$ nach Formel (4) durch einfache Matrixmultiplikationen (wegen der Zentrosymmetrie gibt es nur 30 verschiedene Blöcke). Das L-Gebiet in Fig. 1 entsteht aus dem quadratischen Gebiet in Fig. 2 durch Abtrennung eines Teilquadrates, d. h. durch Hinzunahme der neuen Randbedingungen

$$x_{16} = x_{26} = x_{36} = x_{46} = x_{56} = x_{57} = x_{58} = x_{59} = x_{5,10} = 0.$$

Diese neun Randbedingungen können nach Abschnitt IV gleichzeitig durch eine neunfache rangvermindernde Operation berücksichtigt werden. Wird die Minormatrix von $\mathbf{R}$ (Ordnung 100×9), welche aus den Spalten mit den

Doppelindizes 16, 26, 36, 46, 56, 57, 58, 59, 5,10 besteht, mit **S**, die quadratische nichtsinguläre Matrix 9-ter Ordnung, welche aus den gemeinsamen Elementen von **S** und **S**′ besteht, mit **T** bezeichnet, so erhält man die Inverse $\mathbf{L}^{*(-1)}$ der zum L-Gebiet gehörigen Laplaceschen Operatormatrix aus folgender Formel:

$$\mathbf{L}^{*(-1)} = \mathbf{R} - \mathbf{S}\mathbf{T}^{-1}\mathbf{S}'.$$

Die Berechnung von $\mathbf{T}^{-1}$ erfordert die Inversion einer Matrix 9-ter Ordnung, und man hat — um die oben angegebene Kapazität der Rechenmaschine nicht zu überschreiten — die Matrix **S** in 10 Teilmatrizen (Ordnung 10×9) zu partitionieren.

(Eingegangen am 30. Dezember 1959.)

Literaturverzeichnis

[1] E. Egerváry, On a lemma of Stieltjes on matrices, *Acta Sci. Math. Szeged*, **15** (1953/54), S. 99—103.

[2] E. Egerváry, On hypermatrices whose blocks are commutable in pairs and their application in lattice dynamics, *Acta Sci. Math. Szeged*, **15** (1953/54), S. 211—222.

[3] E. Egerváry, On rank-diminishing operations and their application to the solution of lineare quations, *Zeitschrift für angew. Math. und Phys.*, **11** (1960), S. 376—386.

[4] C. Stéphanos, Sur une extension du calcul des substitutions linéaires, *Journal de mathématiques pures et appliquées*, **6** (1900), S. 73—126.

[5] E. Pascal, *Repertorium der höheren Mathematik* (Leipzig, 1910).

ON SHORTEST NETS WITH MESHES OF EQUAL AREA

By

L. FEJES TÓTH (Budapest)

(*Presented by* G. HAJÓS)

The classical authors of the bee's cell (as PAPPUS THE ALEXANDRINE, KEPLER, RÉAUMUR, MARALDI etc.) considered it as an obvious fact — and also we are inclined to do so — that among the decompositions of the plane into cells of equal area the regular hexagonal tessellation yields the minimal average perimeter of the cells[1]. But it is little known that the proof of this conjecture involves considerable difficulties which have not been surmounted so far. We shall deal with this problem in the Euclidean and non-Euclidean plane under the restriction to tessellations with straight edges. The difficulties inherent in the general case will be illustrated by some examples.

The space underlying our considerations is a "sphere" of curvature K, i. e. the 2-dimensional spherical space ($K>0$), the Euclidean plane ($K=0$) or the hyperbolic plane ($K<0$). Consider on a sphere of curvature K a closed, connected domain D and decompose it by a finite number of simple arcs into a finite number of simply connected subdomains. The subdomains are called faces, or cells, the arcs edges and their endpoints vertices. The vertices decompose the boundary of D into arcs which we shall call, to distinguish them from the inner edges, boundary edges.

Instead of decomposing a domain into cells, we can set up the domain from cells. Therefore we shall find it convenient to speak besides of partitions also of tessellations.

Considering the vertices as at least trivalent atoms and the edges as arms representing the bonds, we may distinguish free and bonded vertices,

[1] See e. g. D'ARCY W. THOMPSON, *On growth and form. II* (Second edition, Cambridge, 1952), where we read on p. 527: "The investigation of the ends of the cell was a more difficult matter than that of its sides, and came later." However, the "problem of the ends" may be considered today as a simple school-problem, while the "problem of the sides" is a question not yet settled in its full generality. We may suppose that the authors mentioned above restricted their attention to congruent cells. — As to the problem of the bee's cell it is a strong restriction to consider first the problem of the sides and then the problem of the ends only for hexagonal prisms (under further restrictions). It would be desirable to consider the whole problem under far more general conditions.

according as only two or at least three edges meet in them (Fig. 1). In view of the following considerations we allow to decompose a boundary edge by free vertices into smaller edges, but cancel the free internal vertices by uniting the two edges issuing from them. In this sense the inner edges of the partition are uniquely given.

If each cell is a polygon, the tessellation is said to be *polygonal.* An inner edge of a polygonal tessellation may be a broken line, but we restrict our attention to *straight edged* polygonal tessellations, each inner edge of which is a segment.

We introduce the *boundary characteristic* c of a polygonal tessellation by

$$c = f - b + 6h - 12,$$

where f and b are the numbers of the free and bonded vertices lying on the boundary of the tessellation and h is the number of the holes, i. e. the

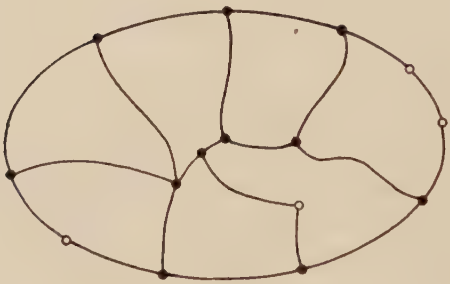


Fig. 1

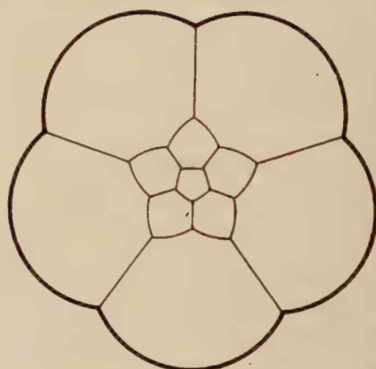


Fig. 2

number of the connected regions of which the complementary of the tessellation consists. If the cells fill up the whole sphere ($K > 0$), we have $h = f = b = 0$, and thus $c = -12$. Later on we shall show that for a tessellation consisting of a finite number of faces of a regular hexagonal tessellation we have $c = 0$. Thus c may be considered as a deviation from this distincted type of tessellations.

Our main result is contained in the following

THEOREM[2]. *Consider on a sphere of curvature K a connected, finite polygonal tessellation of boundary characteristic c. If the tessellation has n*

[2] Special cases of this theorem can be found in my book *Lagerungen in der Ebene, auf der Kugel und im Raum* (Berlin—Göttingen—Heidelberg, 1953, III, § 9, and V, § 9), but without proof of the convexity of the function U for $K > 0$.

cells of equal area t, then the average perimeter l of the cells satisfies the inequality

$$l \geqq U\left(6+\frac{c}{n}, t, K\right)$$

where $U(p, t, K)$ is defined by

$$U(p,t,K)=\begin{cases} \dfrac{2p}{\sqrt{K}} \arccos \dfrac{\cos \dfrac{\pi}{p}}{\cos \dfrac{2\pi-Kt}{2p}} & \text{for} \quad K>0, \\[2ex] 2\sqrt{tp \operatorname{tg} \dfrac{\pi}{p}} & \text{for} \quad K=0, \\[2ex] \dfrac{2p}{\sqrt{-K}} \operatorname{ar\,ch} \dfrac{\cos \dfrac{\pi}{p}}{\cos \dfrac{2\pi-Kt}{2p}} & \text{for} \quad K<0. \end{cases}$$

Equality holds only for tessellations with equal regular cells and at most trihedral vertices.

Observe that, in view of $\operatorname{ar\,ch} z = \frac{1}{i} \arccos z$ and

$$U(p,t,0)=\lim_{K\to 0} U(p,t,K),$$

the function $U(p,t,K)$ may be defined for $K>0$ and $K<0$ by means of one analitical formula and for $K=0$ by postulation of continuity. As to the case of equality note that for tessellations having inner vertices equality cannot hold but in the case of equal regular cells of angles $120°$ (Fig. 2, 3, 4).

Fig. 3

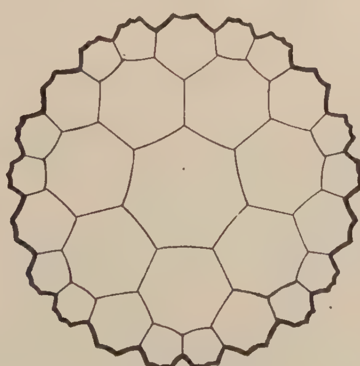


Fig. 4

Our theorem involves the following

COROLLARY. *If U is the union of n distinct cells of a regular tessellation with trihedral vertices, then among the partitions of U into n convex cells of equal area the regular partition has the least total edge-length.*

In order to deduce this from our theorem let us denote the perimeter of a cell of the regular tessellation by $\bar{l}$ and the average perimeter of an irregular partition of U into n convex cells of equal area by l. Furthermore, let $\bar{c}$ and c be the boundary characteristics of the regular and irregular decompositions of U. In view of the convexity of the cells of the irregular decomposition, each bonded vertex of the regular decomposition is, simultaneously, a bonded vertex of the irregular decomposition. Thus $c \leqq c$, whence

$$l > U\left(6+\frac{c}{n}, t, K\right) \geqq U\left(6+\frac{\bar{c}}{n}, t, K\right) = \bar{l}.$$

The basic idea of the proof of our theorem is very simple. Denoting the number of sides of the single cells[3] by $p_1, \ldots, p_n$ and their perimeters by $l_1, \ldots, l_n$ we have, by the isoperimetric property of the regular polygons,

$$l_i \geqq U(p_i, t, K),$$

with equality only if the cell is regular. We shall show, as a simple consequence of Euler's theorem, that

$$p_1 + \cdots + p_n \leqq 6n + c.$$

Here equality holds only for tessellations with at most trihedral vertices. Furthermore, we shall prove that, in the domain of the (p, t, K)-space which is to be considered from the point of view of the problem, $U(p, t, K)$ is a convex, decreasing function of p. These facts involve the desired inequality:

$$nl = l_1 + \cdots + l_n \geqq U(p_1, t, K) + \cdots + U(p_n, t, K) \geqq$$
$$\geqq nU\left(\frac{p_1 + \cdots + p_n}{n}, t, K\right) \geqq nU\left(6+\frac{c}{n}, t, K\right)$$

Equality cannot hold simultaneously in all three inequalities except for the case indicated in our theorem.

In order to prove the inequality $p_1 + \cdots + p_n \leqq 6n + c$, let us remark that we may restrict ourselves to tessellations having not more than trihedral vertices. For, the vertices from which more than three edges emanate can

[3] Note that the vertices of the cells lying on the boundary of a particular cell are considered as to be vertices of this cell too. Thus a cell of a tessellation may have more vertices (sides) than it has effectively as a single polygon.

be "split up" into trihedral vertices and we have only to observe that by this operation $p_1+\cdots+p_n$ increases and c does not decrease. We shall show that, for tessellations of the considered kind, $p_1+\cdots+p_n=6n+c$. This implies our above remark concerning the boundary characteristic of hexagonal tessellations.

Let e be the total number of edges and v the total number of vertices of the tessellation. Then, by Euler's theorem,

$$(n+h)+v=e+2,$$

whence, in view of $3v-f=2e$,

$$p_1+\cdots+p_n=2e-(f+b)=6n+f-b+6h-12=6n+c.$$

Turning now to the proof of the convexity of U as a function of p, we have to consider separately the cases $K=1$, $K=0$ and $K=-1$.

Case $K=1$. We shall show that for $p>2$, $0<t<2\pi$ we have $U_{pp}>0$. With the notations

$$\frac{\pi}{p}=x, \quad \frac{2\pi-t}{2\pi}=\varepsilon, \quad y(x)=\operatorname{arc\,cos}\frac{\cos x}{\cos \varepsilon x},$$

we have $U=2py\left(\frac{\pi}{p}\right)$, whence $p^3U_{pp}=2\pi^2y''\left(\frac{\pi}{p}\right)$. Thus we have only to show that for $0<x<\pi/2$, $0<\varepsilon<1$ the inequality $y''(x)>0$ holds.

On account of

$$\log\cos y=\log\cos x-\log\cos\varepsilon x$$

we obtain

$$-y'\operatorname{tg} y=-\operatorname{tg} x+\varepsilon\operatorname{tg}\varepsilon x,$$

$$-y''\operatorname{tg} y-\frac{y'^2}{\cos^2 y}=-\frac{1}{\cos^2 x}+\frac{\varepsilon^2}{\cos^2\varepsilon x}.$$

Thus the inequality to be proved is

$$y''\operatorname{tg} y=\frac{1}{\cos^2 x}-\frac{\varepsilon^2}{\cos^2\varepsilon x}-\frac{(\operatorname{tg} x-\varepsilon\operatorname{tg}\varepsilon x)^2}{1-\dfrac{\cos^2 x}{\cos^2\varepsilon x}}>0.$$

This reduces by the chain

$$(\cos^2\varepsilon x-\cos^2 x)(\cos^2\varepsilon x-\varepsilon^2\cos^2 x)-\cos^2 x\cos^4\varepsilon x(\operatorname{tg} x-\varepsilon\operatorname{tg}\varepsilon x)^2>0,$$

$$(\sin^2 x-\sin^2\varepsilon x)(\cos^2\varepsilon x-\varepsilon^2\cos^2 x)-$$
$$-(1-\sin^2\varepsilon x)(\sin x\cos\varepsilon x-\varepsilon\cos x\sin\varepsilon x)^2>0,$$

$$-(\sin 2\varepsilon x-\varepsilon\sin 2x)^2+4\sin^2\varepsilon x(\sin x\cos\varepsilon x-\varepsilon\cos x\sin\varepsilon x)^2>0,$$

$$-\sin 2\varepsilon x+\varepsilon\sin 2x+2\sin\varepsilon x(\sin x\cos\varepsilon x-\varepsilon\cos x\sin\varepsilon x)>0$$

of equivalent inequalities to

$$G(x, \varepsilon) = \varepsilon \sin 2x - (1 - \sin x) \sin 2\varepsilon x - 2\varepsilon \sin^2 \varepsilon x \cos x > 0.$$

The proof rests on elementary properties of G as function of ε alone. We have

$$-G_{\varepsilon\varepsilon} = (\cos x - x + x \sin x) \sin 2\varepsilon x + \varepsilon x \cos x \cos 2\varepsilon x,$$

which we write in the form

$$-\frac{2}{\cos x} G_{\varepsilon\varepsilon} = A \sin z + z \cos z, \quad z = 2\varepsilon x, \quad A = 2\frac{\cos x - x + x \sin x}{\cos x}.$$

Writing $x = \frac{\pi}{2} - \xi$, we have for $0 < x < \pi/2$

$$\frac{1}{2} A \cos x = \sin \xi - \left(\frac{\pi}{2} - \xi\right)(1 - \cos \xi) > \xi - \frac{\xi^3}{6} - \left(\frac{\pi}{2} - \xi\right)\frac{\xi^2}{2} =$$

$$= \frac{\xi}{12}(4\xi^2 - 3\pi\xi + 12) > 0,$$

showing for $\varepsilon x \leqq \pi/4$, $0 < \varepsilon < 1$, $0 < x < \pi/2$ the validity of $G_{\varepsilon\varepsilon}(x, \varepsilon) < 0$. This involves, in view of $G(x, 0) = G(x, 1) = 0$, the positivity of $G(x, \varepsilon)$ in the rectangle $0 < \varepsilon < 1$, $0 < x \leqq \pi/4$.

In order to settle the case $\frac{\pi}{4} < x < \frac{\pi}{2}$, let us remark that the equation

$$A \sin z + z \cos z = 0 \qquad (A > 0)$$

has exactly one root for $0 < z < \pi$ and thus at most one root for $0 < z < 2x$. Therefore, $G(x, \varepsilon)$ has, as function of ε, at most one point of inflexion for $0 < \varepsilon < 1$. Thus the inequality $G(x, \varepsilon) > 0$ shall be proved by showing that for $\frac{\pi}{4} < x < \frac{\pi}{2}$ we have

$$-G_\varepsilon(x, 1) = 2x(\cos 2x + \sin x) + 2 \cos x \sin^2 x - \sin 2x > 0,$$

or, what is the same, that for $0 < x < \pi/4$ we have

$$-G_\varepsilon\left(\frac{\pi}{2} - x, 1\right) = 2\left(\frac{\pi}{2} - x\right)(\cos x - \cos 2x) + 2 \sin x \cos^2 x - \sin 2x > 0.$$

But this is true, because

$$-G_\varepsilon\left(\frac{\pi}{2} - x, 1\right) > \frac{\pi}{2}(\cos x - \cos 2x) - (1 - \cos x) \sin 2x >$$

$$> \frac{\pi}{2}\left(\frac{3x^2}{2} - \frac{16x^4}{24}\right) - \frac{x^3}{2} = \frac{x^2}{12}(9\pi - 6x - 4\pi x^2) > 0,$$

and herewith the discussion of the case $K = 1$ is completed.

Case $K=0$. The domain to be considered is now $p \geqq 3$, $t>0$. The condition $U_{pp}>0$ of the convexity in this domain reduces, with the notation $\pi/p=x$, to

$$4x \sin x > 2x - \sin 2x.$$

This is, in view of $2x-\sin 2x < \frac{4x^3}{3}$, certainly satisfied if $\sin x > \frac{x^2}{3}$. But this is true even for $p=\frac{\pi}{x} \geqq 2$.

Case $K=-1$. Now the geometrically possible values of p and t are given by $p \geqq 3$, $0<t<\pi(p-2)$. Introducing the notations

$$\frac{\pi}{p}=x, \quad \frac{2\pi+t}{2\pi}=\varepsilon, \quad y(x)=\operatorname{ar\,ch} \frac{\cos x}{\cos \varepsilon x}$$

we have, analogously as in the case of $K=1$, $p^3 U_{pp}=2\pi^2 y''\left(\frac{\pi}{p}\right)$. Therefore it suffices to show that for $0<x<\pi/2$, $x<\varepsilon x<\pi/2$ we have $y''(x)>0$. By an analogous computation and with the same meaning of $G(x, \varepsilon)$ as above, this inequality turns out to be equivalent to $G(x, \varepsilon)<0$. Recalling that $G(x, \varepsilon)>0$ for $0<\varepsilon<1$, $G(x, 1)=0$, and, for $0<\varepsilon<\pi/2x$, $G(x, \varepsilon)$ has, as function of ε, exactly one point of inflexion, we have for $1<\varepsilon<\pi/2x$, in view of

$$G\left(x, \frac{\pi}{2x}\right)=\frac{\pi}{x} \cos x(\sin x-1)<0,$$

in fact $G(x, \varepsilon)<0$; and this ends the proof of $U_{pp}>0$.

In order to verify that U is a decreasing function of p, it suffices to refer now only to the existence of $\lim\limits_{p \to 0} U(p, t, K)$. This completes the proof of our theorem.

Directing our attention to the general problem, it may be shown that there is a shortest net which decomposes a given domain into a given number of parts of equal area, and it follows (by methods familiar in isoperimetric problems) that the edges must be arcs of constant curvature. The following examples show that in general the edges are not straight segments.

Consider the shortest segment which decomposes an equilateral triangle of side 2 into two parts of equal area. It is easy to show that this segment is parallel to one side, and consequently its length equals $\sqrt{2}=1,41\ldots$. On the other hand, the circular arc centred at a vertex and yielding a similar decomposition has a length equal to $\sqrt{\pi/\sqrt{3}}=1,34\ldots$.

The following interesting remark is due to A. HEPPES. It can be seen easily that from each vertex of the shortest net dividing the sphere $(K=1)$

into $n \geqq 3$ parts of equal area exactly three edges issue and under equal angles. Consider a net with straight edges satisfying this condition. Then all cells have the same number of sides, say, p, and we have $\frac{2\pi}{3}p-\pi(p-2)=\frac{4\pi}{n}$, i. e. $6-p=12/n$. It follows that $n=3$, 4, 6 or 12, on account of which for any other values of $n>2$ the extremal net must have curved edges. This yields a peculiar attraction to the conjecture that for $n=4$, 6 and 12 the solutions are polygonal tessellations and consequently regular ones.[4]

Our last example relates to an analogous problem in Euclidean 3-space. It is known that among the "space fillers" of FEDOROV[5] of equal volume the truncated octahedron has the least surface area. But LORD KELVIN showed that the faces of this polyhedron can be wrapped and its edges curved so as to obtain a new space filler of the same volume and smaller surface area. This fact sets in its true light the intrinsic interest of the conjecture mentioned in the introduction which may now be precised as follows: If U is the union of n distinct faces of a regular hexagonal tessellation, then among all partitions of U into n parts of equal area the regular partition has the least total edge-length.

(Received 30 December 1959)

[4] The case $n=3$ has been settled by HEPPES.

[5] Cf. e. g. pp. 551 552 of the book quoted in footnote [1].

A NON-HAMILTONIAN PLANAR GRAPH

By

W. T. TUTTE (Toronto)

(*Presented by* G. HAJÓS)

A planar graph is made up of *edges* and *vertices*. The vertices are points in a plane. Each edge is a closed arc in the plane joining two vertices and not passing through a third vertex. It is required that no two edges shall intersect, save possibly at a common end. In this paper we suppose the number of edges and vertices to be finite.

A *polygon* in a planar graph G is a simple closed curve made up of edges of the graph. It is a *Hamiltonian* polygon if it passes through all the vertices.

The *valency* of a vertex v is the number of edges having v as an end. We call G *cubic* if the valency of each of its vertices is 3. A cubic planar graph is *cyclically k-connected* if it cannot be broken up into two separate parts, each containing a polygon, by the removal of fewer than k edges.

Hamiltonian polygons of planar graphs, especially cubic ones, are of interest in connection with the Four Colour Conjecture. For it is easily seen that if a planar graph has a Hamiltonian polygon, then the map into which it dissects the plane can be coloured in four colours. Two colours occur alternately on the inside of the polygon and two others on the outside.

P. G. TAIT [1] conjectured that every "true polyhedron", that is every cubic planar graph which is cyclically 3-connected, has a Hamiltonian polygon. A cyclically 3-connected (but not 4-connected) counterexample to this was published by the present author in 1946 [2]. The object of the present paper is to improve upon this result by exhibiting a cyclically 4-connected counterexample to TAIT's conjecture. A brief account of the graph in question was given at the Dobogókő colloquium on graph theory (October 20—22, 1959).

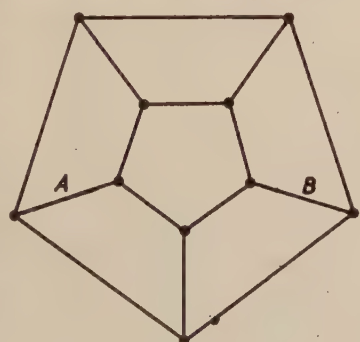


Fig. 1

Let G_1 denote the planar cubic graph shown in Fig. 1. It can be verified by trial that G_1 has no Hamiltonian polygon passing through the two edges A and B.

Now let G be a planar cubic graph having a portion G_2 of the form shown in Fig. 2. Let H be any Hamiltonian polygon of G. Then we have the following theorem:

(1) *If H contains C and D, it contains also E and F. The part of H inside G_2 then consists of two arcs, one containing C and D and the other containing E and F. Moreover, the part of H outside G_2 consists of two arcs, one joining C and E and the other joining D and F.*

PROOF. Suppose H contains C and D. It certainly contains an even number of the edges C, D, E and F. Hence either H contains both E and F or it contains neither of them.

We make the following assumption: Either H does not contain E and F or the part of H inside G_2 consists of two arcs, one joining C and E and the other joining D and F. Clearly, H cannot contain both the edges ab and bc, and it cannot contain both the edges de and ef. Hence H contains both A and B. Let us remove from G_2 the edges C, D, E and F, and contract each of the triangles abc and def to a single vertex. This operation transforms G_2 into G_1, and under our assumption it transforms the part of H inside G_2 into a Hamiltonian polygon of G_1 passing through both A and B, which is impossible.

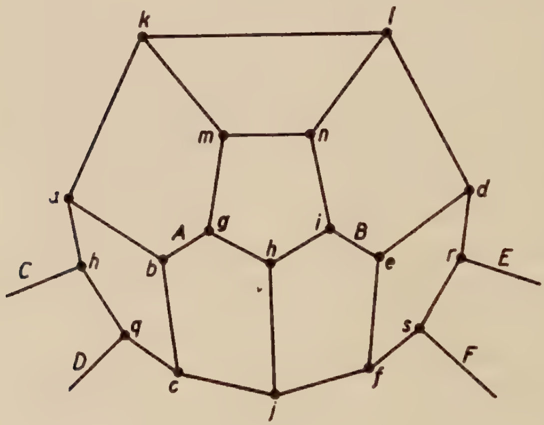


Fig. 2

We deduce that H contains both E and F. The part inside G_2 must consist of two arcs joining C, D, E and F in pairs. We have seen that these pairs are not $\{C, E\}$ and $\{D, F\}$. They are not $\{C, F\}$ and $\{D, E\}$, since the arcs cannot cross. Hence they are $\{C, D\}$ and $\{E, F\}$. Moreover, the part of H outside G_2 must consist of two arcs joining C, D, E and F in pairs. Since H is a single polygon and the arcs cannot cross, these pairs must be $\{C, E\}$ and $\{D, F\}$.

Combining two copies of G_2 we obtain the graph G_3 shown in Fig. 3.

(2) *Let G be a planar cubic graph having G_3 as a part, and let H be a Hamiltonian polygon of G. Then H does not contain all four of the edges P, Q, T and U.*

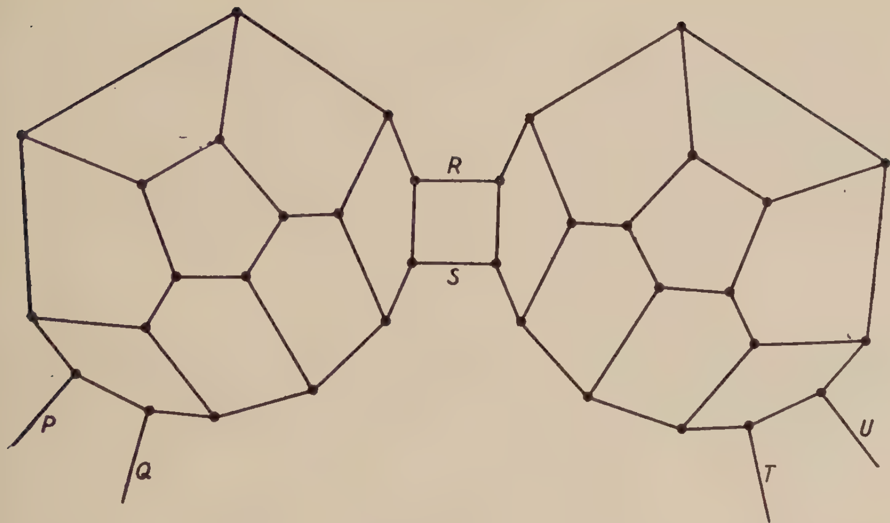


Fig. 3

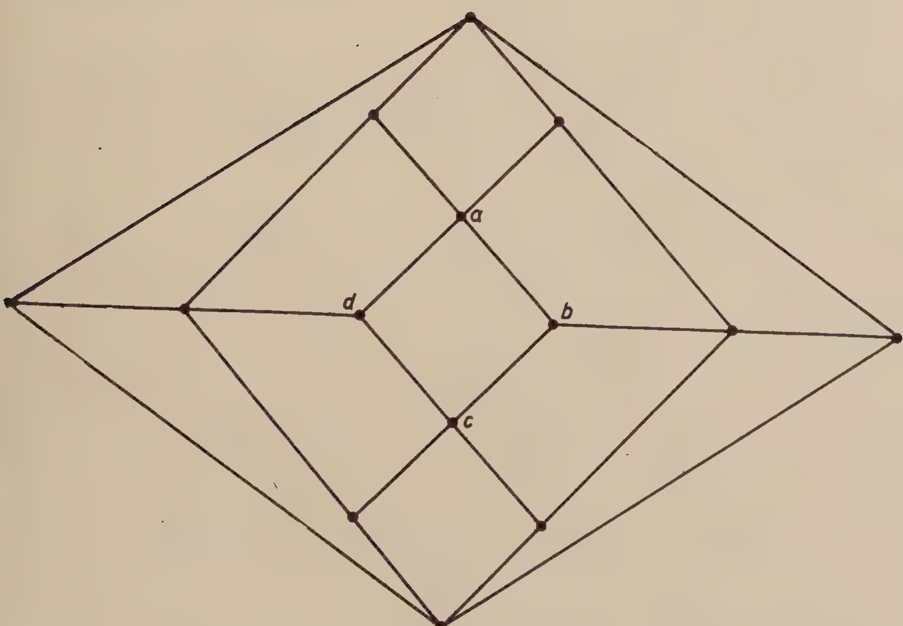


Fig. 4

PROOF. Suppose H does contain these four edges. Then by applying Theorem 1 to either copy of G_2 we find that H contains also R and S. But then Theorem 1 cannot be true for both copies of G_2, since H is a single polygon.

Let us now consider the non-cubic planar graph G_4 of Fig. 4. It has 8 vertices of valency 3 and 6 of valency 4. We form a planar cubic graph G from G_4 by replacing each vertex of valency 4 by a copy of G_3.

(3) *G has no Hamiltonian polygon.*

PROOF. Suppose H is a Hamiltonian polygon of G. Then H passes through the 8 vertices of G_4 of valency 3. It therefore contains 16 of the 24 edges radiating from the 6 copies of G_3. Hence there are two of these copies of G_3 for which H contains all four of the radiating edges. This is contrary to Theorem 2.

It remains to show that G is cyclically 4-connected. Let us define a *Q-subgraph* of G as a connected subgraph K such that each edge of K belongs to some polygon of K. Thus each copy of G_2 or G_3 occurring in G is a Q-subgraph (if we do not count the four radiating edges).

(4) *Let the vertices of a Q-subgraph K of G be partitioned into two disjoint non-null sets V' and V''. Then there are at least two edges of K having an end in each set.*

PROOF. Suppose the theorem false. Then since K is connected, there is just one edge of K joining V' and V''. But this edge cannot belong to any polygon of K, contrary to the definition of a Q-subgraph.

(5) *G is cyclically 4-connected.*

PROOF. Assume the contrary. Then by the removal of at most three edges G can be broken up into two separate parts G' and G'', each containing a polygon. Let U' and U'' be the sets of vertices of G' and G'', respectively. We say that a subgraph K of G is *split* by the pair $\{U', U''\}$ if it includes at least one vertex of each set.

Suppose first that no one of the six copies of G_3 in G is split by $\{U', U''\}$ Consider the Q-subgraph K of G corresponding to the quadrilateral $abcd$ of G_4. It includes two copies of G_3 and two additional vertices b and d. The vertices of G not in K and the edges of G having no end in K define another Q-subgraph K_1 of G. By Theorem 4 the pair $\{U', U''\}$ does not split both K and K_1. For there are at most three edges of G joining U' and U''.

Assume that $\{U', U''\}$ does not split K_1. Then we may suppose that G' is a subgraph of K. But then either G' is joined to G'' by more than

three edges or it has only one vertex b or d. In either case the definition of G' and G'' is contradicted.

We deduce that $\{U', U''\}$ splits K_1 and therefore does not split K. But we may apply a similar argument to each of the Q-subgraphs corresponding to the 11 other quadrilaterals of G_4 equivalent to $abcd$ under the symmetry of G_4. It follows that all the vertices of G belong to the same set U' or U'', which is impossible.

In the remaining case some copy J of G_3 is split by $\{U', U''\}$. The vertices of G not in J and the edges of G having no end in J define another Q-subgraph J_1 of G. By Theorem 4 J_1 is not split by $\{U', U''\}$. We may therefore suppose that G' is a subgraph of J. Now G' is not J, for J is joined to the rest of G by four distinct edges, and similarly G' is not either of the copies of G_2, L_1 and L_2 say, which occur in J. Since $\{U', U''\}$ does not split both the Q-subgraphs L_1 and L_2, it follows readily that G' is a proper subgraph of one of them, let us say of L_1.

Let us consider Fig. 2 as a diagram of L_1. Let R be the Q-subgraph of G defined by k, l, m and n, and R_1 the Q-subgraph defined by the remaining vertices of L_1. Suppose first that the second Q-subgraph is not split by $\{U', U''\}$. Then we may assume G' to be contained in R. This implies $G' = R$, since G' has at least one polygon. But this is impossible, since R is joined to the rest of G by four distinct edges. In the remaining case R is not split, and we may assume G' to be contained in R_1. Since G' has a polygon, it meets each of the arcs $CpabghiedrE$ and $DqcjfsF$. Accordingly, each of these disjoint arcs has two edges joining G' to the rest of G. This contradicts the definition of G'. The Theorem follows.

(Received 31 December 1959)

References

[1] P. G. Tait, Listing's Topologie, *Phil. Mag.* (5), 17 (1884), pp. 30—46; *Scientific Papers*, Vol. II, pp. 85—98.

[2] W. T. Tutte, On Hamiltonian circuits, *Journal London Math. Soc.*, 21 (1946), pp. 98—101.

ÜBER DIE APPROXIMATIONSTHEORETISCHE CHARAKTERISIERUNG GEWISSER FUNKTIONENKLASSEN

Von

D. KRÁLIK (Budapest)

(*Vorgelegt von* G. ALEXITS)

Einleitung

Die approximationstheoretische Untersuchung reeller Funktionen hat eine doppelte Zielsetzung: sie ist einerseits bestrebt, auf Grund der strukturellen Eigenschaften einer Funktion ihre Approximationsgeschwindigkeit mittels einfacher analytischer Ausdrücke (z. B. Polynome oder trigonometrischer Polynome) festzustellen („direktes Problem"), andererseits sucht sie aus der Größenordnung der Approximation auf die strukturellen Eigenschaften der approximierten Funktion zu schließen („umgekehrtes oder inverses Problem"). Falls man für eine Funktionenklasse das direkte wie auch das umgekehrte Problem mit derselben Approximationsgrößenordnung beantworten kann, erhält man die vollständige approximationstheoretische Charakterisierung dieser Funktionenklasse. Im Fall der stetigen Funktionen hat D. JACKSON [7], [8] das direkte Problem gelöst, indem er zeigte, daß die beste Approximation einer Funktion mittels Polynome bzw. trigonometrischer Polynome dieselbe Größenordnung hat wie ihr Stetigkeitsmodul. Die Ergebnisse von JACKSON können leicht auf die Funktionen des allgemeineren Funktionenraumes L^p [1] ausgedehnt werden: bedeuten $\omega_p(\delta, f)$ den Stetigkeitsmodul der Funktion $f(x)$ in der Metrik von L^p, [2] $E_n^{(p)}(f)$ ihre beste Approximation mittels trigono-

[1] $f(x) \in L^p = L^p[0, 2\pi]$ $(p \geqq 1)$, wenn $|f(x)|^p$ in $[0, 2\pi]$ im Lebesgueschen Sinne integrierbar und $f(x)$ nach 2π periodisch ist. $\|f\|_p = \left(\int_0^{2\pi} |f(x)|^p \, dx\right)^{1/p}$ ist die Norm von $f(x)$. $L^\infty = L^\infty[0, 2\pi]$ soll den Raum $C[0, 2\pi]$ der stetigen Funktionen bedeuten. Im folgenden haben wir es immer mit Funktionen zu tun, die im Intervall $[0, 2\pi]$ definiert und nach 2π periodisch sind.

[2] $$\omega_p(\delta, f) = \omega_p(\delta, f; 0, 2\pi) = \sup_{|h| \leqq \delta} \left(\int_0^{2\pi} |f(x+h) - f(x)|^p \, dx\right)^{1/p} \quad (1 \leqq p < \infty),$$

$$\omega_\infty(\delta, f) = \omega(\delta, f) = \sup_{|h| \leqq \delta} \max_{\substack{x \\ x+h}\} \in [0, 2\pi]} |f(x+h) - f(x)|$$

ist der gewöhnliche Stetigkeitsmodul.

metrischer Polynome höchstens n-ter Ordnung in derselben Metrik, so i

(1) $$E_n^{(p)}(f)=O\left[\omega_p\left(\frac{1}{n},f\right)\right].$$

Die Beantwortung des umgekehrten Problems im Fall der Klass Lip α $(0<\alpha<1)$ stammt von S. BERNSTEIN [5], wodurch sich zusammen m dem Ergebnis von JACKSON die vollständige approximationstheoretische Cha rakterisierung dieser Funktionenklasse ergab. Es wurde später dieses Result auf die allgemeinere Klasse Lip (α, p)[3] $(0<\alpha<1; p\geqq 1)$ ausgedehnt.[4] Di Lösung des Problems für $\alpha=1$ hat G. ALEXITS [1] und unabhängig von ihr M. ZAMANSKY [13] gefunden. Hinsichtlich des umgekehrten Problems ist D LA VALLÉE POUSSIN [12] zu noch allgemeineren Resultaten als BERNSTEI gelangt.

In dieser Arbeit wollen wir das umgekehrte Problem für eine recht all gemeine Funktionenklasse vollständig lösen. Auf Grund der Jacksonsche Lösung des direkten Problems erhalten wir also die approximationstheore tische Charakterisierung der betreffenden Funktionenklassen. Die Charakteri sierung der Klasse Lip (α, p) $(1\leqq p\leqq+\infty)$ und vieler anderen klassische Funktionenklassen (z. B. der Dini—Lipschitzschen Klassen α-ter Ordnung) er gibt sich als Spezialfall unserer Resultate.

Wir führen die folgende Funktion ein: Sei $\lambda(x)$ eine für positive x Werte definierte positive, monoton abnehmende Funktion, für die $\lim\limits_{x\to+\infty}\lambda(x)=$ ist. Es soll weiterhin eine geeignete Zahl $0<\gamma<1$ existieren, für welch $\lambda(x)x^\gamma$ bei genügend großen x monoton nicht abnehmend ist.[5] Wir sager die Funktion $f(x)$ gehöre zur Funktionenklasse L_λ^p, wenn $f(x)\in L^p$ $(p\geqq 1)$ un

(2) $$\omega_p(\delta,f)=O\left[\lambda\left(\frac{1}{\delta}\right)\right]$$

ist. Wir beweisen zunächst den folgenden

SATZ 1. *Es ist* $f(x)\in L_\lambda^p$ $(1\leqq p\leqq+\infty)$ *genau dann, wenn die Beziehun*

(3) $$E_n^{(p)}(f)=O[\lambda(n)]$$

[3] Eine Funktion $f(x)\in L^p$ gehört zur Klasse Lip (α, p), wenn $\omega_p(\delta,f)=O(\delta^\alpha)$ is
[4] Siehe z. B. [2], S. 226.
[5] In der zitierten Arbeit von DE LA VALLÉE POUSSIN wird eine ähnliche Funktion $\Omega($ eingeführt, welche der einschränkenden Bedingung $\int^{+\infty}\Omega(x)\frac{dx}{x}<\infty$ zu genügen hat. Ob wohl diese Arbeit von DE LA VALLÉE POUSSIN mit der unseren eine gewisse Ähnlichkeit ha schließt diese Einschränkung die Charakterisierung wichtiger Funktionenklassen, wie z. der Dini—Lipschitzschen α-ter Ordnung, aus.

besteht, wobei (abgesehen von der Konstanten) als beste approximierende trigonometrische Polynome die Fejérschen Mittel der Fourierreihe von $f(x)$ betrachtet werden können.

Es ist zu bemerken, daß der Inhalt des Satzes 1 im Spezialfall des Raumes $C(=L^{\infty})$ schon von STEČKIN [11] bewiesen wurde, ohne jedoch die Rolle der Fejérschen Mittel erkannt zu haben.

Es ist bekannt der Satz von PRIVALOFF [9], nach welchem eine Funktion $f(x)$ zugleich mit ihrer konjugierten $\tilde{f}(x)$ für $0<\alpha<1$ zur selben Klasse $\mathrm{Lip}\,(\alpha, p)$ gehören. Wenn aber $\alpha=1$ und $p=1$ bzw. $p=+\infty$ sind, so gilt dies nicht mehr, und wir können von der Struktur der konjugierten Funktion auch im Fall anderer Funktionenklassen nur weniges behaupten. Wir haben auch in dieser Richtung Resultate erzielt, die denen von PRIVALOFF analog sind. Es besteht nämlich der folgende

SATZ 2. *Ist $f\in L^p_\lambda$ und $1<p<+\infty$, so ist auch $\tilde{f}\in L^p_\lambda$. In den Fällen $p=1$ und $p=+\infty$ können wir das nur behaupten, wenn die Funktion $\lambda(x)$ außer den schon erwähnten Eigenschaften noch die folgende*[6] *besitzt:*

$$\int_{1/\delta}^{+\infty} \frac{\lambda(x)}{x}\,dx = O\left[\lambda\left(\frac{1}{\delta}\right)\right]. \tag{4}$$

Die Bedingung (4) wird z. B. erfüllt, wenn für eine geeignete positive Zahl $\beta\leqq\gamma$ die Funktion $\lambda(x)x^{\beta}$ für genügend große x-Werte monoton nicht zunimmt. Wählt man $\lambda(x)=x^{-\alpha}$ $(0<\alpha<1)$, so ergibt sich der Satz von PRIVALOFF als Spezialfall unseres Satzes 2. Auf Grund des Satzes 2 können wir dann die im Satz 1 ausgesprochene Charakterisierung der Klasse L^p_λ folgenderweise aussprechen:

SATZ 3. *Damit $f\in L^p_\lambda$ und $\tilde{f}\in L^p_\lambda$, ist das Bestehen der beiden Beziehungen*[7]

$$\|\sigma_n(x,f)-f(x)\|_p=O[\lambda(n)],\quad \|\tilde{\sigma}_n(x,f)-\tilde{f}(x)\|_p=O[\lambda(n)] \tag{5}$$

notwendig, und das Bestehen einer von diesen hinreichend.

Die Lokalisation der auf trigonometrische Polynome bezogenen Bernsteinschen Ungleichung [4] ermöglicht uns die Beantwortung des umgekehrten Problems für unsere Klasse L^p_λ in lokalisierter Form auszusprechen. Wir haben nämlich den folgenden

6 Siehe diesbezüglich auch [3], wo einige Berührungspunkte mit unserem Satz 2 für den Fall $p=+\infty$ zu finden sind.

7 $\sigma_n(x,f)$ bzw. $\tilde{\sigma}_n(x,f)$ bedeuten das n-te Fejérsche Mittel der Fourierreihe bzw. der konjugierten Reihe von $f(x)$.

SATZ 4. *Ist die Funktion $f(x)$ im Intervall $[a, b] \subset [0, 2\pi]$ L^p-integrierbar und ihre beste Approximation mittels trigonometrischer Polynome höchstens n-ter Ordnung im Teilintervall $[a, b]$ von der Größenordnung*

$$E_n^{(p)}(f; a, b) = \min_{T_n} \left(\int_a^b |f(x) - T_n(x)|^p dx \right)^{1/p} = O[\lambda(n)],^8 \tag{6}$$

so hat der Stetigkeitsmodul $\omega_p(\delta, f; a_1, b_1)$[9] für ein beliebiges, völlig im Inneren von $[a, b]$ liegendes Teilintervall $[a_1, b_1]$ die Größenordnung $O\left[\lambda\left(\frac{1}{\delta}\right)\right]$, wo die Konstante in O von a_1 und b_1 abhängt.

Was das direkte Problem anbetrifft, werden wir für die lokale Approximation in der Metrik der Räume $C[a, b]$ und $L[a, b]$ völlig befriedigende Resultate erhalten. Im Fall der Räume L^p $(1 < p < +\infty)$ ist uns aber nicht gelungen, ein diesen gleichwertiges Resultat zu erreichen. Unsere diesbezüglichen Ergebnisse sind die folgenden:

SATZ 5. *Hat eine in $[0, 2\pi]$ definierte periodische Funktion $f(x)$ im Intervall $[a, b] \subset [0, 2\pi]$ den Stetigkeitsmodul*

$$\omega_p(\delta, f; a, b) = O\left[\lambda\left(\frac{1}{\delta}\right)\right], \tag{7}$$

so gilt im Fall $p = +\infty$

$$E_n(f; a, b) = O[\lambda(n)], \tag{8}$$

und im Fall $p = 1$

$$E_n^{(p)}(f; a_1, b_1) = O[\lambda(n)] \tag{9}$$

in einem beliebigen Teilintervall $[a_1, b_1]$ von $[a, b]$, wo die Konstante in O von a_1 und b_1 abhängen kann. In den Fällen $1 < p < +\infty$ gilt (9) für $p\gamma \leqq 1$, wo γ die in der Definition von $\lambda(x)$ eingeführte Konstante bedeutet.[10]

[8] Im Fall $p = +\infty$ haben wir $E_n^{(\infty)}(f; a, b) = E_n(f; a, b) = \min\limits_{T_n} \operatorname{Max}\limits_{x \in [a, b]} |f(x) - T_n(x)|$.

[9] $\omega_p(\delta, f; a_1, b_1) = \sup\limits_{|h| \leqq \delta} \left(\int_{a_1}^{b_1} |f(x+h) - f(x)|^p dx \right)^{1/p}$, wo $\delta > 0$ so klein sein soll, daß $x + h \in [a, b]$ ist.

$\omega_\infty(\delta, f; a_1, b_1) = \omega(\delta, f; a_1, b_1) = \sup\limits_{|h| \leqq \delta} \max\limits_{x \text{ und } x+h \in [a_1, b_1]} |f(x+h) - f(x)|$.

[10] Die Lokalisationsfrage der Approximation im Raume C wurde unter anderen Gesichtspunkten auch von FREY [6] und anderen untersucht.

Beweis der Sätze 1 und 2

Beim Beweis des umgekehrten Teiles des Satzes 1 spielt die folgende, von STEČKIN [11] bewiesene[11] Beziehung

$$\omega_p\left(\frac{1}{n}, f\right) \leqq \frac{C}{n} \sum_{\nu=1}^{n} E_\nu^{(p)}(f) \tag{10}$$

die entscheidende Rolle, wo C eine Konstante ist. Aus (3) und (10) folgt nämlich auf Grund der angenommenen Eigenschaften von $\lambda(x)$

$$\omega_p\left(\frac{1}{n}, f\right) = \frac{O(1)}{n} \sum_{\nu=1}^{n} \lambda(\nu) = O\left(\frac{1}{n}\right) \sum_{\nu=1}^{n} \lambda(\nu)\nu^{\gamma}\nu^{-\gamma} = O(n^{-1+\gamma})\lambda(n) \sum_{\nu=1}^{n} \nu^{-\gamma}.$$

Da die letzte Summe $O(n^{1-\gamma})$ ist, erhalten wir mithin

$$\omega_p\left(\frac{1}{n}, f\right) = O[\lambda(n)],$$

woraus sich (2) auf Grund der Eigenschaften von $\lambda(x)$ und des monotonen Charakters von $\omega_p(\delta, f)$ ergibt.

Der direkte Teil des Satzes 1 ist nach (1) evident. Wir haben nur noch die erste Beziehung von (5) zu beweisen.[12] Aus der Gültigkeit von (2) und aus der Formel

$$\sigma_n(x, f) - f(x) = \frac{1}{\pi(n+1)} \int_0^{\pi/2} [f(x+2t) + f(x-2t) - 2f(x)] \frac{\sin^2(n+1)t}{\sin^2 t}\, dt$$

[11] In der Originalarbeit von STECKIN ist (10) nur für den gewöhnlichen Stetigkeitsmodul $\omega\left(\frac{1}{n}, f\right)$ und für die Approximation $E_n(f)$ in der Metrik von $L^\infty[0, 2\pi]$ ausgesprochen und bewiesen. Sein Beweis kann aber wörtlich auf den allgemeinen Fall von $L^p[0, 2\pi]$ übertragen werden, da die auf trigonometrische Polynome bezogene Bernsteinsche Ungleichung, weiterhin die im Beweis vorkommenden anderen Beziehungen in entsprechender Weise auch für den Raum L^p gültig sind. Eine genauere Rechnung zeigt die Gültigkeit der Beziehung

$$\omega_p\left(\frac{1}{n}, f\right) \leqq \frac{C}{n} \sum_{\nu=1}^{n} E_\nu^{(p)}(f) + \frac{\|f\|_p}{n},$$

wo C eine absolute Konstante ist. Das zweite Glied auf der rechten Seite ist aber für das Weitere belanglos. Das unnötige Glied $\frac{\|f\|_p}{n}$ kann dadurch beseitigt werden, daß man nur diejenigen Funktionen $\varphi \in L^p$ betrachtet, für welche $\|\varphi\|_p \leqq 1$ ist. Im Fall $\|f\|_p > 1$ betrachten wir also die Funktion $\varphi(x) = \frac{f(x)}{\|f\|_p}$. Da $\omega_p(\delta, \varphi) = \frac{1}{\|f\|_p}\omega_p(\delta, f)$ ist, haben $\varphi(x)$ und $f(x)$ dieselbe Stetigkeitsstruktur.

[12] Das hat für den Fall $L^\infty[0, 2\pi]$ schon ZYGMUND getan (s. [14], I, S. 91). Unser Beweis bezieht sich auf den Fall L^p mit $1 \leqq p < +\infty$.

erhalten wir auf Grund der Minkowskischen Ungleichung[13]

$$\|\sigma_n(x,f)-f(x)\|_p \leqq \frac{1}{\pi(n+1)}\int_0^{\pi/2}\|f(x+2t)+f(x-2t)-2f(x)\|_p \frac{\sin^2(n+1)t}{\sin^2 t}dt=$$

$$=O(1)\frac{1}{\pi(n+1)}\int_0^{\pi/2}\lambda\left(\frac{1}{t}\right)\frac{\sin^2(n+1)t}{t^2}dt.$$

Zerlegen wir das Integral in die Teile $\int_0^{1/n}$ und $\int_{1/n}^{\pi/2}$. Die Größenordnung des ersten Teiles ist wegen des monoton abnehmenden Charakters von $\lambda(x)$

$$O(n)\int_0^{1/n}\lambda\left(\frac{1}{t}\right)dt=O[\lambda(n)].$$

Die Auswertung des zweiten Integrals ergibt auf Grund der Eigenschaften von $\lambda(x)$

$$O(n^{-1})\int_{1/n}^{\pi/2}\lambda\left(\frac{1}{t}\right)\left(\frac{1}{t}\right)^{\gamma}t^{\gamma-2}dt=O(n^{-1})\lambda(n)n^{\gamma}\int_{1/n}^{\pi/2}t^{\gamma-2}dt=O[\lambda(n)].$$

Damit ist der Beweis des Satzes 1 zu Ende.

Was den Satz 2 anbetrifft, gilt im Fall $1<p<+\infty$ auf Grund der bekannten von M. RIESZ [10] herrührenden Beziehung

$$\|\tilde{f}(x+h)-\tilde{f}(x)\|_p \leqq A_p\|f(x+h)-f(x)\|_p \leqq A_p\omega_p(\delta,f) \quad (|h|\leqq\delta),$$

das heißt

$$\omega_p(\delta,\tilde{f}) \leqq A_p\omega_p(\delta,f),$$

wo die Konstante A_p nur von p abhängt. In den Fällen $p=1$ und $p=+\infty$ gilt die Rieszsche Ungleichung nicht, und wir haben die Integralformel

$$\tilde{f}(x+h)-\tilde{f}(x)=-\frac{1}{\pi}\left\{\int_{-\pi}^{\pi}\frac{f(x+t)-f(x+h)}{2\operatorname{tg}\frac{t-h}{2}}dt-\int_{-\pi}^{\pi}\frac{f(x+t)-f(x)}{2\operatorname{tg}\frac{t}{2}}dt\right\}$$

abzuschätzen.[14] Wir werden die Abschätzung für die Funktionenklasse L_λ^∞ ausführen. (Für L_λ^1 geht die Abschätzung ganz analog, nur haben wir statt der Beziehung $|f(x+h)-f(x)|=O\left[\lambda\left(\frac{1}{h}\right)\right]$ die Integralbeziehung $\int_0^{2\pi}|f(x+h)-f(x)|dx=O\left[\lambda\left(\frac{1}{h}\right)\right]$ zu benützen.) Zerlegen wir die Integrale

[13] Siehe z. B. [15], S. 67.

n die Teilintegrale: $\int\limits_{-\pi}^{-2h} + \int\limits_{-2h}^{2h} + \int\limits_{2h}^{\pi}$. Für die Größenordnung des ersten und ritten Teiles erhalten wir auf der gewöhnlichen Weise[14] $O\left[\lambda\left(\frac{1}{h}\right)\right]$. Für das weite Teilintegral ergibt sich aber die Größenordnung $O\left[\lambda\left(\frac{1}{h}\right)\right]$ nur unter uhilfenahme von (4). Es ist nämlich

$$\int\limits_{-2h}^{2h} \frac{f(x+t)-f(x+h)}{2\,\mathrm{tg}\,\frac{t-h}{2}}\,dt + \int\limits_{-2h}^{2h} \frac{f(x+t)-f(x)}{2\,\mathrm{tg}\,\frac{t}{2}}\,dt =$$

$$= O(1) \int\limits_{-2h}^{2h} \frac{\lambda\left(\frac{1}{|t-h|}\right)}{|t-h|}\,dt + O(1) \int\limits_{-2h}^{2h} \frac{\lambda\left(\frac{1}{|t|}\right)}{|t|}\,dt = I_1 + I_2.$$

urch die Substitution $t-h=\tau$ ergibt sich für I_1 bei Beachtung von (4)

$$O(1) \int\limits_{0}^{3h} \frac{\lambda\left(\frac{1}{\tau}\right)}{\tau}\,d\tau = O(1) \int\limits_{1/3h}^{+\infty} \frac{\lambda(x)}{x}\,dx = O\left[\lambda\left(\frac{1}{3h}\right)\right] = O\left[\lambda\left(\frac{1}{h}\right)\right].$$

anz analog erhalten wir für I_2 dieselbe Größenordnung. Damit haben wir en Satz 2 bewiesen.

Beweis der Sätze 4 und 5

Der Beweis des Satzes 4 verläuft ähnlich zu dem des Satzes 1, nur aben wir statt der Beziehung (10) die entsprechende „lokalisierte" Beziehung

$$\omega_p\left(\frac{1}{n}, f; a_1, b_1\right) \leqq \frac{C(a_1, b_1)}{n} \sum_{\nu=1}^{n} E_\nu^{(p)}(f; a, b)\,^{15}$$

[14] Vgl. [15], S. 156—157, oder [14], I, S. 121—122.

[15] Die genauere Rechnung ergibt

$$\omega_p\left(\frac{1}{n}, f; a_1, b_1\right) \leqq \frac{C_1(a_1, b_1)}{n} \sum_{\nu=1}^{n} E_\nu^{(p)}(f; a, b) + \frac{C_2(a_1, b_1)}{n} \|f\|_p,$$

o jetzt $\|f\|_p = \left(\int\limits_a^b |f(x)|^p\,dx\right)^{1/p}$ ist. Das zweite Glied auf der rechten Seite kann aber bei r Abschätzung von $\omega_p\left(\frac{1}{n}, f; a_1, b_1\right)$ gegenüber dem ersten vernachlässigt werden (s. ch [11]).

anzuwenden. Der Beweis dieser Ungleichung ist analog dem von (10), nur hat man die gewöhnliche, auf das ganze Intervall $[0, 2\pi]$ bezogene Bernsteinsche Ungleichung durch ihre lokalisierte Form zu ersetzen ([4]).

Wir führen den Beweis des Satzes 5 zuerst für den Fall $L^{\infty}[a, b] = C[a, b]$ durch. Betrachten wir die folgende Hilfsfunktion: Setzen wir $\varrho = \dfrac{2\pi - b}{3}$ und

$$g(x) = \begin{cases} f(x) & x \in [a, b], \\ f(a) & x \in [0, a], \\ f(b) & x \in [b, b+\varrho], \\ f(a) & x \in [b+2\varrho, 2\pi], \end{cases}$$

im Intervall $[b+\varrho, b+2\varrho]$ sei $g(x)$ linear und $g(x+2\pi) = g(x)$. (Wäre $b = 2\pi$, so sei $g(x) = f(x)$ in $[a, b]$, $g(x) = f(a)$ in $\left[\dfrac{2a}{3}, a\right]$, $g(x) = f(b) = f(2\pi)$ in $\left[0, \dfrac{a}{3}\right]$ und in $\left[\dfrac{a}{3}, \dfrac{2a}{3}\right]$ sei $g(x)$ linear.) Gilt (7) mit $p = +\infty$, so gilt auch $\omega(\delta, g; 0, 2\pi) = O\left[\lambda\left(\dfrac{1}{\delta}\right)\right]$. Die auf die Funktion $g(x)$ bezogene Relation (1) ergibt dann

$$E_n(g; 0, 2\pi) = O[\lambda(n)],$$

und da $E_n(f; a, b) \leqq E_n(g; 0, 2\pi)$, erhalten wir im Fall $p = +\infty$ die behauptete Beziehung (8).

Im Fall L^p $(1 \leqq p < +\infty)$ betrachten wir die Hilfsfunktion

$$g(x) = \begin{cases} f(x) & x \in [a_1, b_1], \\ 0 & x \notin [a_1, b_1], \end{cases}$$

wo $a_1 < b_1$ derartige Punkte zwischen a und b bezeichnen, in welchen

$$\int_{a_1}^{a_1 \pm h} |f(t)|^p dt = O(h), \quad \int_{b_1}^{b_1 \pm h} |f(t)|^p dt = O(h) \tag{11}$$

ist. Diese Beziehungen gelten in jedem Punkte a_1 und b_1, wo $|f(x)|^p$ die Ableitung ihres Integrals ist, d. h. in $[a, b]$ fast überall. Man kann also a_1 und b_1 an a und b so nahe wählen, daß sich die Länge des Intervalls $[a_1, b_1]$ von $b-a$ um beliebig wenig unterscheidet. Wir wollen das Integral

$$I(h) = \int_0^{2\pi} |g(x+h) - g(x)|^p dx$$

durch die Zerlegung

$$I(h)=\int_0^{a_1-h}+\int_{a_1-h}^{a_1}+\int_{a_1}^{b_1-h}+\int_{b_1-h}^{b_1}+\int_{b_1}^{2\pi}\quad(h>0)$$

abschätzen. Das erste und das letzte Integral verschwindet, da in diesen Intervallen $g(x+h)\equiv g(x)$ ist. In $[a_1-h, a_1]$ ist $g(x)\equiv 0$, also ist der Wert des zweiten Integrals nach (11)

$$\int_{a_1-h}^{a_1}|g(x+h)|^p dx=\int_{a_1}^{a_1+h}|f(t)|^p dt=O(h),$$

und die gleiche Größenordnung ergibt sich auch für das vierte Integral. Wir erhalten ferner für das dritte Integral

$$\int_{a_1}^{b_1-h}|g(x+h)-g(x)|^p dx=\int_{a_1}^{b_1-h}|f(x+h)-f(x)|^p dx\leqq[\omega_p(\delta,f;a,b)]^p.$$

Aus (7) ergibt sich somit im Fall $p=1$ die Abschätzung $I(h)=O\left[\lambda\left(\frac{1}{\delta}\right)\right]$, und daher

$$\omega_1(\delta,g;0,2\pi)=\sup_{|h|\leqq\delta}I(h)=O\left[\lambda\left(\frac{1}{\delta}\right)\right].$$

Ein dem Fall $p=+\infty$ analoger Gedankengang ergibt die Beziehung (9). Für $p>1$ können wir ähnliches nur behaupten, wenn $\lambda^p\left(\frac{1}{\delta}\right)\geqq C\delta$ ist, wo C eine positive Konstante ist, was für $p\gamma\leqq 1$ wegen der Eigenschaften von $\lambda(x)$ sicher gilt.

(Eingegangen am 17. Februar 1960.)

Literaturverzeichnis

[1] G. Alexits, Sur l'ordre de grandeur de l'approximation d'une fonction périodique par les sommes de Fejér, *Acta Math. Acad. Sci. Hung.*, 3 (1952), S. 29—40.

[2] Н. И. Ахиезер, Лекции по теории аппроксимации (Москва—Ленинград, 1947).

[3] Н. К. Бари, О наилучщей приближении тригонометрическими полиномами двух сопряженных функций, Изв. Акад. Наук СССР, Сер. мат., 19 (1955), S. 285—302.

[4] Н. К. Бари, Обобщение неравенств С. Н. Бернштейна и А. А. Маркова, ДАН СССР, 90 (1953), S. 701—703.

[5] S. N. Bernstein, Sur l'ordre de la meilleure approximation des fonctions par des polynômes de degré donné, *Mémoires couronnées Acad. Roy. de Belgique* (2), 4 (1912), S. 1—104.

[6] T. Frey, A legjobb polinomapproximáció lokalizálásáról, *MTA Mat. és Fiz. Oszt. Közl.*, **7** (1957), S. 403—412 (ungarisch).

[7] D. Jackson, *Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen*, Dissertation (Göttingen, 1911).

[8] D. Jackson, On approximation by trigonometric sums and polynomials, *Trans. Amer. Math. Soc.*, **14** (1912), S. 491—515.

[9] J. Privaloff, Sur les fonctions conjuguées, *Bull. Soc. Math. de France*, **44** (1916), S. 100—103.

[10] M. Riesz, Sur les fonctions conjuguées, *Mat. Zeitschrift*, **27** (1928), S. 218—244.

[11] С. Б. Стечкин, О порядке наилучших приближений непрерывных функций, Изв. Акад. Наук СССР, Сер. мат., **15** (1951), S. 234—235.

[12] de la Vallée Poussin, *Leçons sur l'approximation des fonctions d'une variable réelle* (Paris, 1919), S. 53—62.

[13] M. Zamansky, Classes de saturation de certains procédés d'approximation des séries de Fourier des fonctions et applications à quelques problèmes d'approximation, *Ann. Sci. École Norm. Sup.* (3), **66** (1949), S. 19—93.

[14] A. Zygmund, *Trigonometric series. I—II* (Cambridge, 1959).

[15] A. Zygmund, *Trigonometrical series* (Warsawa—Lwow, 1935).

ÜBER APPROXIMATIONEN MIT DEN ARITHMETISCHEN MITTELN ALLGEMEINER ORTHOGONALREIHEN

Von

G. ALEXITS (Budapest), Mitglied der Akademie, und D. KRÁLIK (Budapest)

Prof. P. TURÁN *anläßlich seines 50. Geburtstages zugeeignet*

Einleitung

Abschätzungen über die Güte der Annäherung einer Funktion $f(x)$ mit Linearkombinationen vorgegebener Funktionen $f_1(x), f_2(x), \ldots$ sind nur für spezielle Funktionensysteme $\{f_n(x)\}$ bekannt. Insbesondere gibt es klassische Ergebnisse im Fall, daß $\{f_n(x)\}$ das System der trigonometrischen oder algebraischen Polynome ist. Dagegen scheint das Problem der Güte der Approximation mit den Mitteln einer allgemeinen Orthogonalentwicklung

$$(1) \qquad f(x) \sim \sum_{n=0}^{\infty} c_n \varphi_n(x)$$

überhaupt nicht untersucht zu sein. Der Grund dafür steckt im Umstand, daß die Orthogonalreihen auf die strukturellen Eigenschaften der entwickelten Funktion $f(x)$ im allgemeinen nicht reagieren (vgl. BANACH [5]). Nur in neuester Zeit haben in dieser Richtung J. MEDER [11] und K. TANDORI [13] einige Resultate erzielt. Das Ergebnis von TANDORI, welches das von MEDER beträchtlich erweitert, lautet folgenderweise:

Ist $\{\lambda(n)\}$ eine positive, monoton gegen Unendlich konvergente Zahlenfolge, welche der Bedingung

$$\lambda(n^2) = O[\lambda(n)]$$

genügt und ist auch die Bedingung

$$\sum_{n=2}^{\infty} c_n^2 (\log\log n)^2 \lambda^2(n) < \infty$$

erfüllt, so gelten für die $(C, 1)$-Mittel $\sigma_n(x)$ der Orthogonalreihe (1) fast überall die Beziehungen:

$$\sigma_n(x) - f(x) = o_x\left(\frac{1}{\lambda(n)}\right),$$

$$\lambda(n)|\sigma_n(x) - f(x)| \leqq F(x) \qquad (n = 1, 2, \ldots, F \in L^2),$$

wo $f(x)$ die durch den Riesz—Fischerschen Satz bis auf eine Nullmenge eindeutig bestimmte Funktion bedeutet. Sowohl der Beweis von MEDER als auch

der von TANDORI beruhen auf bekannten Methoden der Theorie der Orthogonalreihen.

Wir können bedeutend weitergehende Resultate erzielen, wenn wir uns der Frage von einer ganz anderen Seite nähern. In der Theorie der Fourierreihen konnte man nämlich mittels reihentheoretischer Betrachtungen ganz allgemeiner Natur verschiedene Approximationssätze beweisen; es ist sogar dadurch gelungen, wichtige Funktionenklassen mit den Methoden der konstruktiven Funktionentheorie zu charakterisieren (ALEXITS [1], [2] und KRÁLIK [9]). Wenden wir diese Methode auf das obige Problem an, so ergeben sich Sätze, welche einerseits das Ergebnis von TANDORI weit verallgemeinern, andererseits auch zur Behandlung anderer Probleme den Weg bahnen, u. a. zu einer Verallgemeinerung des Weylschen Begriffes des Integrals gebrochener Ordnung.

Im § 1 beweisen wir ein allgemeines reihentheoretisches Lemma, dessen Anwendung im § 2 die Erweiterung des Approximationssatzes von TANDORI, und einen Approximationssatz für die $(C, 1)$-Mittel gewisser Orthogonalpolynomreihen ergeben wird, bei welchem auch die Struktur der entwickelten Funktion zur Geltung kommt. Im § 3 verallgemeinern wir, ebenfalls aus unserem reihentheoretischen Hauptlemma ausgehend, den Begriff des Weylschen Integrals gebrochener Ordnung wie auch die diesbezüglichen tiefliegenden Ergebnisse von HARDY und LITTLEWOOD [8].

§ 1. Das reihentheoretische Hauptlemma

Bezeichne $\Lambda(x)$ eine für $x>0$ definierte positive, von unten konvexe, nichtwachsende Funktion mit der Eigenschaft: es gibt eine feste Zahl $0<\gamma<1$ derart, daß $\Lambda(x)x^\gamma$ für genügend große x monoton nichtabnehmend ist. Sei ferner $\Theta(x)$ eine Funktion mit ähnlichen Eigenschaften wie $\Lambda(x)$, nur soll in der Monotonieforderung für $\Theta(x)x^\eta$ statt γ eine Zahl $0<\eta<1$ mit $\gamma+\eta<1$ stehen.

Sei nun R ein Banachscher Raum, $u_1, u_2, \ldots$ Elemente von R und σ_n das n-te $(C, 1)$-Mittel[1] der Reihe

$$\sum_{n=1}^{\infty} u_n.$$

HAUPTLEMMA. *Konvergieren die σ_n (in der Metrik von R) gegen ein Element $\sigma \in R$ mit dem Annäherungsgrad*

$$\|\sigma-\sigma_n\| = O[\Lambda(n)], \tag{2}$$

[1] $\sigma_n = \sum_{k=1}^{n}\left(1-\frac{k}{n+1}\right) u_k$ und entsprechende Formeln gelten auch für die weiteren Reihen.

so approximieren die $(C,1)$*-Mittel* $\sigma_n(\theta)$ *der Reihe* $\sum_{n=1}^{\infty} u_n\theta(n)$ *(in der Metrik von R) ein Element* $\sigma(\theta)\in R$ *mit dem Annäherungsgrad*

$$\|\sigma(\theta)-\sigma_n(\theta)\|=O[\Lambda(n)\theta(n)]. \tag{3}$$

Bezeichnen wir das n-te $(C,1)$-Mittel der Reihe $(u_1-\sigma)+u_2+u_3+\cdots$ mit σ_n^*, das der Reihe $\theta(1)(u_1-\sigma)+\theta(2)u_2+\theta(3)u_3+\cdots$ mit $\sigma_n^*(\theta)$, so ergeben sich aus der Annahme (2) wegen $\sigma_n^*=\sigma_n-\sigma+\frac{\sigma}{n+1}$ die folgenden Beziehungen:

$$\|\sigma_n^*\|\leqq\|\sigma_n-\sigma\|+\frac{\|\sigma\|}{n+1}=O[\Lambda(n)], \tag{4}$$

$$\|\sigma_n(\theta)-\sigma_m(\theta)\|\leqq\|\sigma_n^*(\theta)-\sigma_m^*(\theta)\|+\frac{\theta(1)\|\sigma\|}{n} \qquad (m>n). \tag{5}$$

Zwischen den Mitteln $\sigma_n^*(\theta)$ und σ_n^* besteht die Beziehung[2]

$$\sigma_n^*(\theta)=\sum_{k=1}^{n-1}\left(1-\frac{k}{n+1}\right)(k+1)\sigma_k^*\Delta^2\theta(k)+$$
$$+\frac{2}{n+1}\sum_{k=1}^{n-1}(k+1)\sigma_k^*\Delta\theta(k+1)+\theta(n)\sigma_n^*,$$

so daß die folgende Ungleichung gilt:

$$\begin{aligned}\|\sigma_n^*(\theta)-\sigma_m^*(\theta)\|\leqq&\left\|\sum_{k=1}^{n-1}\left(1-\frac{k}{n+1}\right)(k+1)\sigma_k^*\Delta^2\theta(k)-\right.\\ &\left.-\sum_{k=1}^{m-1}\left(1-\frac{k}{m+1}\right)(k+1)\sigma_k^*\Delta^2\theta(k)\right\|+\frac{2}{n+1}\sum_{k=1}^{n-1}(k+1)\|\sigma_k^*\|\Delta\theta(k+1)+\\ &+\frac{2}{m+1}\sum_{k=1}^{m-1}(k+1)\|\sigma_k^*\|\Delta\theta(k+1)+\theta(n)\|\sigma_n^*\|+\theta(m)\|\sigma_m^*\|.\end{aligned} \tag{6}$$

Für $m>n$ ist $\left(1-\frac{k}{m+1}\right)-\left(1-\frac{k}{n+1}\right)<\frac{k}{n}$ und $1-\frac{k}{m+1}<1$, daher erhalten wir nach (4) für das erste Glied der rechten Seite von (6) die Abschätzung

$$O\left(\frac{1}{n}\right)\sum_{k=1}^{n-1}k^2\Lambda(k)\Delta^2\theta(k)+O(1)\sum_{k=n}^{m-1}k\Lambda(k)\Delta^2\theta(k)=S_1+S_2.$$

Die Eigenschaften von $\Lambda(x)$ und $\theta(x)$ ergeben nach einer Abelschen Um-

[2] Vgl. [12].

formung:

$$S_1 = O\left(\frac{1}{n}\right)\sum_{k=1}^{n-1} k^{2-\gamma} \Lambda(k) k^{\gamma} \Delta^2\theta(k) = O(n^{-1+\gamma})\Lambda(n)\sum_{k=1}^{n-1} k^{2-\gamma}\Delta^2\theta(k) =$$

$$= O(n^{-1+\gamma})\Lambda(n)\left\{\sum_{k=1}^{n} k^{1-\gamma}\Delta\theta(k) + \Delta\theta(n)\sum_{k=1}^{n} k^{1-\gamma}\right\}.$$

Aus der Monotonie von $\theta(x)x^{\eta}$ folgt, daß auch $\log\theta(x)x^{\eta}$ nichtabnehmend ist, also eine nichtnegative Ableitung hat. Das bedeutet so viel wie

$$-\theta'(\xi) \leqq \frac{\theta(\xi)}{\xi}\eta.$$

Ist ξ ein entsprechend gewählter Punkt zwischen k und $k+1$, so folgt daher

$$\Delta\theta(k) = -\theta'(\xi) \leqq \frac{\theta(\xi)}{\xi}\eta < \frac{\theta(k)}{k}. \tag{7}$$

Daraus ergibt sich bei Beachtung der Monotonie von $\theta(x)x^{\eta}$ und $\gamma+\eta<1$

$$S_1 = O(n^{-1+\gamma})\Lambda(n)\left\{\sum_{k=1}^{n}\frac{k^{\eta}\theta(k)}{k^{\gamma+\eta}} + \theta(n)n^{1-\gamma}\right\} = O[\Lambda(n)\theta(n)].$$

Für die Abschätzung von S_2 erhalten wir auf ähnliche Weise

$$S_2 = O(1)\sum_{k=n}^{m-1} k\Lambda(k)\Delta^2\theta(k) = O[\Lambda(n)]\sum_{k=n}^{m-1} k\Delta^2\theta(k) =$$

$$= O[\Lambda(n)]\left\{\sum_{k=n}^{m}\Delta\theta(k) + m\Delta\theta(m) + (n-1)\Delta\theta(n)\right\} =$$

$$= O[\Lambda(n)]\{\theta(n)+\theta(m)\} = O[\Lambda(n)\theta(n)].$$

Damit ist die Größenordnung $O[\Lambda(n)\theta(n)]$ für das erste Glied der rechten Seite von (6) hergeleitet. Das zweite und dritte Glied haben dieselbe Struktur, es genügt also nur das eine von ihnen, z. B. das zweite abzuschätzen. Wir erhalten dafür bei Beachtung von (7) die Abschätzung

$$O\left(\frac{1}{n}\right)\sum_{k=1}^{n-1} k^{1-\gamma}\Lambda(k)k^{\gamma}\Delta\theta(k) = O(n^{-1+\gamma})\Lambda(n)\sum_{k=1}^{n-1} k^{1-\gamma}\Delta\theta(k) =$$

$$= O(n^{-1+\gamma})\Lambda(n)\sum_{k=1}^{n-1} k^{-\gamma}\theta(k) = O[\Lambda(n)\theta(n)].$$

Es bleibt nur mehr die Abschätzung der beiden letzten Glieder von (6) übrig. Diese haben aber offensichtlich die Größenordnung $O[\Lambda(n)\theta(n)]$. Damit ist die Ungleichung

$$\|\sigma_n^*(\theta) - \sigma_m^*(\theta)\| = O[\Lambda(n)\theta(n)]$$

für jedes $m > n$ bewiesen. Hieraus ergibt sich nach (5) bei Beachtung der Eigenschaften von $\Lambda(x)$ und $\theta(x)$ die erwünschte Beziehung (3). Wir bemerken, daß in unserem soeben bewiesenen Hauptlemma auch $\Lambda(x) \equiv$ Konst. und $\theta(x) \equiv$ Konst. angenommen werden darf, da wir im Beweis nur solche Eigenschaften der Konvexität ausgenützt haben, welche den $\Lambda(x)$ und $\theta(x)$ auch in diesen Grenzfällen zukommen. Wir bemerken ferner, daß im Beweis nur die Konvexität von $\theta(x)$, nicht aber die von $\Lambda(x)$ ausgenützt wurde. Wir haben die Konvexität von $\Lambda(x)$ nur deshalb gefordert, weil wir im folgenden davon ausgiebig Gebrauch machen werden.

§ 2. Approximationssätze für die $(C, 1)$-Mittel allgemeiner Orthogonalreihen und gewisser Orthogonalpolynomentwicklungen

Als erste Anwendung unseres Hauptlemmas beweisen wir den folgenden

SATZ 1. *Sei $\{\varphi_n(x)\}$ ein im Intervall $[a, b]$ definiertes, beliebiges Orthonormalsystem und $\{c_n\}$ eine reelle Zahlenfolge mit $\sum_{n=0}^{\infty} c_n^2 < \infty$. Nehmen wir an, daß die Bedingung*

$$\sum_{n=1}^{\infty} c_n^2 \vartheta^2(n) < \infty \tag{8}$$

die $(C, 1)$-Summierbarkeit der Reihe

$$\sum_{n=0}^{\infty} c_n \varphi_n(x) \tag{9}$$

auf einer Menge $E \subset [a, b]$ sichert, wo $\vartheta(x)$ eine positive, monoton gegen $+\infty$ wachsende Funktion ist. Bedeutet $\lambda(x)$ eine positive, von unten konkave, monoton gegen $+\infty$ strebende Funktion, für welche $\frac{x^\gamma}{\lambda(x)}$ mit festem $0 < \gamma < 1$ bei genügend großen x monoton nicht abnimmt, so folgt aus dem Erfülltsein der Bedingung

$$\sum_{n=1}^{\infty} c_n^2 \lambda^2(n) \vartheta^2(n) < \infty,$$

daß die $(C, 1)$-Mittel $\sigma_n(x)$ der Orthogonalreihe (9) auf E mit dem Annäherungsgrad

$$|\sigma_n(x) - f(x)| = o_x\left(\frac{1}{\lambda(n)}\right) \tag{10}$$

approximieren, wo $f(x) \in L^2[a, b]$ die durch den Riesz—Fischerschen Satz der Reihe (9) zugeordnete Funktion bedeutet.

Der Beweis dieses Satzes folgt unmittelbar aus unserem Hauptlemma. Es gibt nämlich eine geeignete positive, genügend langsam gegen $+\infty$ strebende, von unten konkave Funktion $\nu(x)$, so daß $\frac{x^\omega}{\lambda(x)\nu(x)}$ für ein geeignetes $0<\gamma\leqq\omega<1$ monoton nicht abnimmt und außerdem

$$\sum_{n=1}^{\infty} c_n^2\lambda^2(n)\nu^2(n)\vartheta^2(n)<\infty$$

gilt. Aus (8) folgt also nach Annahme, daß die $(C,1)$-Mittel der Reihe

$$\sum_{n=1}^{\infty} c_n\lambda(n)\nu(n)\varphi_n(x) \tag{11}$$

auf der Menge E konvergieren, sie sind mithin auf E von der Größenordnung $O_x(1)$. Da die Funktionen $\Lambda(x)\equiv$Konst. und $\theta(x)=\frac{1}{\lambda(x)\nu(x)}$ allen Forderungen unseres Hauptlemmas genügen, folgt nach diesem, daß die $(C,1)$-Mittel $\sigma_n(x)$ der Reihe (9) in den Punkten der Menge E der Beziehung

$$|\sigma_n(x)-f(x)|=O_x\left[\frac{1}{\lambda(n)\nu(n)}\right]=o_x\left[\frac{1}{\lambda(n)}\right]$$

genügen, womit unsere Behauptung bewiesen ist.

Im Spezialfall $\vartheta(x)=\log\log x$ konvergieren die $(C,1)$-Mittel der Reihe (11) nach bekannten Sätzen der Theorie der Orthogonalreihen in $[a,b]$ fast überall, und sie bleiben dem Absolutbetrag nach unter einer festen Funktion $F(x)\in L^2[a,b]$. Daraus folgt — wie es aus dem Beweisgang unseres Hauptlemmas leicht ersichtlich ist — für alle n und für fast alle x die Beziehung:

$$\lambda(n)\nu(n)|\sigma_n(x)-f(x)|\leqq\psi(x),$$

wo $\psi(x)\in L^2[a,b]$, womit wir den Approximationssatz von TANDORI als Spezialfall unseres Satzes 1 bewiesen haben. Er ist aber in unserer Fassung auch recht erweitert, da wir uns von der bei TANDORI auftretenden stark einschränkenden Bedingung $\lambda(n^2)=O[\lambda(n)]$ entledigt haben.[3]

Im vorangehenden Satz spielte die Struktur der entwickelten Funktion $f(x)$ sozusagen keine Rolle. Ist aber $\{\varphi_n(x)\}$ ein im Intervall $[-1,+1]$ durch eine Gewichtsfunktion

$$0\leqq\varrho(x)\leqq\frac{\text{Konst.}}{\sqrt{1-x^2}} \tag{12}$$

[3] Die Bedingung $\lambda(n^2)=O[\lambda(n)]$ bedeutet im wesentlichen so viel, daß $\lambda(n)$ ungefähr von der Größenordnung $O(\log^r n)$ mit $r>0$ ist.

eindeutig bestimmtes Orthonormalpolynomsystem $\{p_n(x)\}$, so kommt die Struktur der nach $\{p_n(x)\}$ entwickelten stetigen Funktion

$$f(x) \sim \sum_{n=0}^{\infty} c_n p_n(x) \quad \left(c_n = \int_{-1}^{+1} f(t)\varrho(t)p_n(t)dt\right) \tag{13}$$

bei der Approximation durch Vermittlung der Koeffizienten c_n zur Geltung. Es gilt nämlich der folgende

SATZ 2. *Erfüllt der Stetigkeitsmodul $\omega(\delta, f)$ der stetigen Funktion $f(x)$ in $[-1, 1]$ die Bedingung*

$$\omega(\delta, f) = O\left[\frac{1}{\sqrt{\Phi(1/\delta)}}\right]$$

mit positivem, monoton gegen $+\infty$ strebendem $\Phi(x)$, welches der Bedingung

$$\int_{1}^{+\infty} \frac{\lambda'(x)\lambda(x)}{\Phi(x)} dx < +\infty$$

genügt, wo $\lambda(x)$ die im Satz 1 definierte Funktion ist, gilt ferner für die Gewichtsfunktion $\varrho(x)$ außer (12) *auch die Ungleichung*

$$\varrho(x) \geqq \varrho_0 > 0 \tag{14}$$

in einem völlig im Inneren von $(-1, 1)$ liegenden Teilintervall $[\alpha, \beta] \subset (-1, 1)$, so gilt für die $(C, 1)$-Mittel $\sigma_n(x)$ der Reihe (13) *in (α, β) fast überall die Beziehung* (10).

Mit der Transformation $x = \cos\xi$ übergeht $f(x)$ in die Funktion $f^*(\xi) = f(\cos\xi)$, für deren Stetigkeitsmodul im Intervall $[0, \pi]$ die Beziehung

$$\omega(\delta, f^*) \leqq \omega(\delta, f) = O\left[\frac{1}{\sqrt{\Phi(1/\delta)}}\right]$$

besteht. Auf Grund eines Zusammenhanges zwischen Struktur- und Koeffizientenbedingung[4] konvergiert die Reihe $\sum_{n=1}^{\infty} a_n^2 \lambda^2(n)$, wo $a_n = \frac{2}{\pi}\int_0^{\pi} f^*(\xi)\cos n\xi d\xi$ bedeutet. Nach einem Ergebnis von ALEXITS [3] konvergiert dann wegen (12) auch die Reihe $\sum_{n=1}^{\infty} c_n^2 \lambda^2(n)$. Wie im Beweis des Satzes 1 gibt es auch hier eine geeignete Funktion $\nu(x)$ mit den dortigen Eigenschaften, so daß auch

[4] Vgl. [4].

die Reihe $\sum_{n=1}^{\infty} c_n^2 \lambda^2(n) \nu^2(n)$ konvergiert. Nach einem Satz von G. FREUD [7] folgt somit aus (14) die Konvergenz der $(C, 1)$-Mittel der Reihe $\sum_{n=1}^{\infty} c_n \lambda(n) \nu(n) p_n(x)$ in (α, β) fast überall. Dieselbe Schlußweise wie im Beweis des Satzes 1 ergibt nun für fast alle $x \in (\alpha, \beta)$ die Relation (10).

§ 3. Verallgemeinerung des Weylschen Integrals gebrochener Ordnung und der Sätze von Hardy und Littlewood

Das Weylsche Integral α-ter Ordnung $(\alpha > 0)$ der nach 2π periodischen Funktion $f(x) \in L[0, 2\pi]$ lautet:[5]

$$f_\alpha(x) = \frac{1}{\Gamma(\alpha)} \int_{-\infty}^{x} f(t)(x-t)^{\alpha-1} dt,$$

wo $f(x)$ der Bedingung

$$\int_0^{2\pi} f(x) dx = 0$$

genügt; $f_\alpha(x)$ kann auch als die Faltung von $f(x)$ und $\Psi_\alpha(x)$ in der Form

$$f_\alpha(x) = \frac{1}{\pi} \int_0^{2\pi} f(t) \Psi_\alpha(x-t) dt$$

geschrieben werden, wo

$$\Psi_\alpha(x) = \sum_{n=1}^{\infty} \left(\cos \frac{\pi\alpha}{2} \frac{\cos nx}{n^\alpha} + \sin \frac{\pi\alpha}{2} \frac{\sin nx}{n^\alpha} \right)$$

ist. Die Funktion $f_\alpha(x)$ kann auch durch ihre Fourierreihe

$$f_\alpha(x) = \cos \frac{\pi\alpha}{2} \sum_{n=1}^{\infty} \frac{A_n(x)}{n^\alpha} + \sin \frac{\pi\alpha}{2} \sum_{n=1}^{\infty} \frac{B_n(x)}{n^\alpha} \tag{15}$$

definiert werden, wo $A_n(x) = a_n \cos nx + b_n \sin nx$ das n-te Glied der Fourierreihe von $f(x)$ und $B_n(x) = a_n \sin nx - b_n \cos nx$ das n-te Glied der konjugierten Reihe bedeutet. Wir werden nun durch eine der Formel (15) ähnliche Beziehung eine Funktion einführen, die als eine Verallgemeinerung des Integrals $f_\alpha(x)$ betrachtet werden kann. Das tun wir in wenigen Schritten.

[5] Vgl. [14], Bd. II, S. 133—142.

Sei $\Lambda(x)$ die im § 1 definierte Funktion, welche aber außer den dort angegebenen Eigenschaften auch die folgende besitzt:

$$\int_1^{\infty} \frac{\Lambda(x)}{x} dx < \infty. \tag{16}$$

Es ist wohlbekannt,[6] daß die Reihe $\sum_{n=1}^{\infty} \Lambda(n) \cos nx$ (mit eventueller Ausnahme der Stellen $x \equiv 0 \pmod{2\pi}$) überall gegen eine Funktion $\varphi(x) \in L[0, 2\pi]$ konvergiert, und sie ist die Fourierreihe von $\varphi(x)$. Dasselbe können wir von der Reihe $\sum_{n=1}^{\infty} \Lambda(n) \sin nx$ nur dann behaupten, wenn auch die Bedingung

$$\sum_{n=1}^{\infty} \Delta\Lambda(n) \log n < \infty \tag{17}$$

erfüllt ist.[7] Man sieht sofort, daß die Bedingung (17) mit $\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} < +\infty$ und folglich mit (16) gleichwertig ist.[8] Erfüllt also $\Lambda(x)$ die im § 1 gemachten Voraussetzungen samt der Bedingung (16), so konvergiert die Reihe

$$C_1(\Lambda) \sum_{n=1}^{\infty} \Lambda(n) \cos nx + C_2(\Lambda) \sum_{n=1}^{\infty} \Lambda(n) \sin nx$$

(mit eventueller Ausnahme der Punkte $x \equiv 0 \pmod{2\pi}$) überall gegen eine L-integrierbare Funktion $\Psi_\Lambda(x)$, und sie ist die Fourierreihe von $\Psi_\Lambda(x)$. Die Bestimmung der Konstanten $C_1(\Lambda)$ und $C_2(\Lambda)$ ist hier belanglos. Bedeutet $f(x) \in L[0, 2\pi]$ eine beliebige, nach 2π periodische Funktion mit der Fourierreihe

$$\sum_{n=1}^{\infty} A_n(x) \equiv \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

so existiert nach dem Faltungssatz[9] die Funktion

$$f_\Lambda(x) = \frac{1}{\pi} \int_0^{2\pi} f(t) \Psi_\Lambda(x-t) dt$$

für fast alle $x \in [0, 2\pi]$, sie ist im Intervall $[0, 2\pi]$ L-integrierbar, nach 2π periodisch und

$$C_1(\Lambda) \sum_{n=1}^{\infty} \Lambda(n) A_n(x) + C_2(\Lambda) \sum_{n=1}^{\infty} \Lambda(n) B_n(x) \tag{18}$$

6 Vgl. [14], Bd. I, S. 183.
7 Vgl. [14], Bd. I, S. 185—186.
8 Vgl. [14], Bd. I, S. 185—186.
9 Vgl. [14], Bd. I, S. 36.

ist ihre Fourierreihe. Wir wollen $f_\Lambda(x)$ das Λ-Integral der Funktion $f(x)$ nennen und durch die Reihe (18) definieren. Wählen wir speziell $\Lambda(x) = x^{-\alpha}$ $(\alpha > 0)$, $C_1(\Lambda) = \cos\frac{\pi\alpha}{2}$, $C_2(\Lambda) = \sin\frac{\pi\alpha}{2}$, so erhalten wir $f_\Lambda(x) = f_\alpha(x)$ und $\Psi_\Lambda(x) = \Psi_\alpha(x)$, d. h. wir finden in diesem Spezialfall das Weylsche Integral α-ter Ordnung wieder.

Bezeichnet $\theta(x)$ die im § 1 definierte Funktion, welche aber außer den dort angegebenen Eigenschaften auch der Bedingung (16) genügt, so können wir auch von einem θ-Integral $f_\theta(x)$ von $f(x)$ sprechen, das genau so definiert wird wie $f_\Lambda(x)$. Wählen wir nun die sonst belanglosen Konstanten $C_1(\Lambda)$, $C_2(\Lambda)$, $C_1(\theta)$, $C_2(\theta)$ derart, daß sie den Bedingungen

$$C_1(\Lambda\theta) = C_1(\Lambda)C_1(\theta) - C_2(\Lambda)C_2(\theta),$$
$$C_2(\Lambda\theta) = C_1(\Lambda)C_2(\theta) + C_1(\theta)C_2(\Lambda)$$

genügen, so erhalten wir nach leichter Rechnung

$$(f_\Lambda)_\theta = (f_\theta)_\Lambda = f_{\Lambda\theta},$$

wo $\Lambda\theta = \Lambda(x)\theta(x)$ ist. Setzen wir speziell $\Lambda(x) = x^{-\alpha}$, $\theta(x) = x^{-\beta}$ $(\alpha > 0, \beta > 0)$, so ist $\Lambda(x)\theta(x) = x^{-\alpha-\beta}$, und wir haben die bekannten Beziehungen zwischen den Integralen gebrochener Ordnung vor uns.[10]

Wir beweisen jetzt einige Sätze über Λ-Integrale, welche sich unmittelbar aus unserem reihentheoretischen Hauptlemma ergeben, und die bekannten auf die Integrale gebrochener Ordnung bezogenen Ergebnisse von HARDY und LITTLEWOOD [8] als Spezialfälle enthalten.

SATZ 3. *Ist $f(x) \in L[0, 2\pi]$, $f(x+2\pi) = f(x)$, so gilt für fast alle $x \in [0, 2\pi]$ die Beziehung*

$$f_\Lambda(x) - \sigma_n(x; f_\Lambda) = O_x[\Lambda(n)],$$

wo $\sigma_n(x; f_\Lambda)$ das n-te (C, 1)-Mittel der Fourierreihe von $f_\Lambda(x)$ bedeutet.

[10] Vgl. [14], Bd. II, S. 134. — Bei der Wahl der Konstanten $C_1(\Lambda)$, $C_2(\Lambda)$ haben wir eine sehr große Freiheit, weshalb die Definition des Λ-Integrals durchaus nicht eindeutig ist. Dieser Umstand ist aber kein Nachteil, da hiedurch die Anwendungsmöglichkeiten des Λ-Integrals größer werden. Wollte man jedoch die Konstanten $C_1(\Lambda)$, $C_2(\Lambda)$ derart festsetzen, daß dadurch die Definition des Λ-Integrals eindeutig wird und im Fall $\Lambda(x) = x^{-\alpha}$ gerade das Weylsche Integral ergibt, so könnte man z. B. die folgende Definition von $C_1(\Lambda)$, $C_2(\Lambda)$ annehmen: Bezeichne α das Infimum aller $\gamma > 0$, für welche $x^\gamma \Lambda(x)$ monoton nichtabnehmend ist. Dann setze man $C_1(\Lambda) = \cos\frac{\pi\alpha}{2}$, $C_2(\Lambda) = \sin\frac{\pi\alpha}{2}$. Man sieht leicht, daß die im Text getroffenen Bedingungen für die Konstanten C_1, C_2 für alle in den Anwendungen wichtigen Funktionen Λ und θ erfüllt sind.

Die $(C, 1)$-Mittel $\sigma_n(x; f)$ bzw. $\tilde{\sigma}_n(x; f)$ der Fourierreihe bzw. der konjugierten Reihe von $f(x)$ konvergieren nämlich fast überall, sie sind also fast überall von der Größenordnung $O_x(1)$, woraus auf Grund des Hauptlemmas und der Definition (18) von $f_A(x)$ unsere Behauptung folgt.

SATZ 4. *Sind $f(x) \in L[0, 2\pi]$ und $\tilde{f}(x) \in L[0, 2\pi]$,*[11] *so haben wir $f_A(x) \in L^1_A$.*[12]

Auf Grund unserer Voraussetzungen sind nämlich die $(C, 1)$-Mittel der Reihen $\sum_{n=1}^{\infty} A_n(x)$ bzw. $\sum_{n=1}^{\infty} B_n(x)$ in der Metrik von $L[0, 2\pi]$ beschränkt. Nach dem Hauptlemma konvergieren also die $(C, 1)$-Mittel $\sigma_n(x; f_A)$ der Reihe (18) in der Metrik von $L[0, 2\pi]$ gegen $f_A(x)$ mit der Approximationsgeschwindigkeit

$$\int_0^{2\pi} |\sigma_n(x; f_A) - f_A(x)|\, dx = O[\Lambda(n)],$$

woraus schon unsere Behauptung folgt (vgl. [10], Satz 1 und 3).

SATZ 5. *Ist $f(x) \in L^p[0, 2\pi]$, $1 < p \leqq +\infty$, so gilt $f_A(x) \in L^p_A$.*

Die konjugierte Funktion $\tilde{f}(x)$ gehört nämlich in diesem Fall ebenfalls zur Klasse $L^p[0, 2\pi]$, und die $(C, 1)$-Mittel der Reihen $\sum_{n=1}^{\infty} A_n(x)$ bzw. $\sum_{n=1}^{\infty} B_n(x)$ sind in der Metrik von $L^p[0, 2\pi]$ beschränkt. Der übrige Teil des Beweises verläuft ebenso wie der des Satzes 4. Wir können jetzt sogar noch mehr behaupten, da die Absolutbeträge von $\sigma_n(x; f)$ und $\tilde{\sigma}_n(x; f)$ nach einem bekannten Satz von HARDY und LITTLEWOOD[13] unter je einer L^p-integrierbaren Funktion bleiben:

$$|\sigma_n(x; f)| \leqq g_1(x), \quad |\tilde{\sigma}_n(x; f)| \leqq g_2(x)$$

mit $g_1 \in L^p$, $g_2 \in L^p$. Die Funktionen $g_1(x)$ und $g_2(x)$ sind fast überall endlich, aus unserem Hauptlemma folgt also, daß die $\sigma_n(x; f_A)$ in fast allen Punkten

[11] $\tilde{f}(x)$ bedeutet die konjugierte Funktion von $f(x)$.

[12] Die Funktionenklasse L^p_A $(1 \leqq p \leqq +\infty)$ besteht aus denjenigen, in $[0, 2\pi]$ definierten, nach 2π periodischen Funktionen, die samt ihrer p-ten Potenz im Lebesgueschen Sinne integrierbar sind und deren in der Metrik des Raumes $L^p[0, 2\pi]$ genommener Stetigkeitsmodul

$$\omega_p(\delta, f) = \sup_{|h| \leqq \delta} \left(\int_0^{2\pi} |f(x+h) - f(x)|^p\, dx \right)^{1/p}$$

der Bedingung $\omega_p(\delta, f) = O\left[\Lambda\left(\frac{1}{\delta}\right)\right]$ genügt. Vgl. diesbezüglich [10].

[13] Vgl. [14], Bd. I, S. 155—156.

gegen $f_\Lambda(x)$ konvergieren und in diesen Punkten gilt

$$|\sigma_n(x;f_\Lambda)-f_\Lambda(x)|\leqq g(x)\Lambda(n),$$

wo $g(x)=\text{Konst.}\max\{g_1(x),g_2(x)\}$ ist. Setzen wir $g(x)=+\infty$ in der Nullmenge, in welcher $\{\sigma_n(x;f_\Lambda)\}$ eventuell nicht konvergiert, so erhalten wir den folgenden

SATZ 6. *Ist $f\in L^p$ $(1<p<+\infty)$, so besteht die Ungleichung*

$$|\sigma_n(x;f_\Lambda)-f_\Lambda(x)|\leqq g(x)\Lambda(n)$$

für alle Werte von x mit $g\in L^p$.

SATZ 7. *Gehört die Funktion $f(x)$ zur Funktionenklasse L_Λ^p $(1<p<+\infty)$, so gilt für ihr θ-Integral*

$$\left(\int_0^{2\pi}|f_\theta(x)-\sigma_n(x;f_\theta)|^p dx\right)^{1/p}=\|f_\theta(x)-\sigma_n(x;f_\theta)\|_p=O[\Lambda(n)\theta(n)],$$

mithin gehört $f_\theta(x)$ zur Funktionenklasse $L_{\Lambda\theta}^p$. — Im Fall $p=1$ und $p=+\infty$ bestehen die Beziehungen

$$\int_0^{2\pi}|f_\theta(x)-\sigma_n(x;f_\theta)|dx=O[\Lambda(n)\theta(n)]$$

bzw.

$$\max_{0\leqq x\leqq 2\pi}|f_\theta(x)-\sigma_n(x;f_\theta)|=O[\Lambda(n)\theta(n)]$$

unter der zusätzlichen Annahme

$$\int_n^{+\infty}\frac{\Lambda(x)}{x}dx=O[\Lambda(n)].$$

Unsere Behauptungen sind unmittelbare Korollare unseres Hauptlemmas und der Sätze 1, 2 bzw. 3 in der Arbeit [10] (für den Fall $p=+\infty$ vgl. auch N. BARY [6]).

Wir wiederholen, daß sich die Sätze 3—7 im Spezialfall $\Lambda(x)=x^{-\alpha}$, $\theta(x)=x^{-\beta}$ mit $\alpha>0$, $\beta>0$ und $\alpha+\beta<1$ auf die mehrfach erwähnten Hardy—Littlewoodschen Sätze reduzieren, also sind sie tatsächlich Verallgemeinerungen von diesen. Das Interessante daran ist, daß unsere Verallgemeinerungen aus einfachen approximationstheoretischen Betrachtungen leicht folgen, während der Originalbeweis des Spezialfalles von HARDY und LITTLEWOOD zwar direkt, aber recht langwierig ist; eine dem obigen Gedanken ähnliche Vereinfachung hat einer von uns schon früher erzielt (vgl. [9]).

(Eingegangen am 14. März 1960.)

Literaturverzeichnis

[1] G. Alexits, Sur l'ordre de grandeur de l'approximation d'une fonction par les moyennes de sa série de Fourier, *Mat. és Fiz. Lapok*, **48** (1941), S. 410—421 (ungarisch), S. 421—422 (französisch).

[2] G. Alexits, Sur l'ordre de grandeur de l'approximation d'une fonction périodique par les sommes de Fejér, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 29—42.

[3] G. Alexits, Sur la convergence et la sommabilité presque partout des séries de polynômes orthogonaux, *Acta Sci. Math. Szeged,* **12** B (1950), S. 223—225.

[4] G. Alexits und D. Králik, Über die Bedeutung der strukturellen Eigenschaften einer Funktion für die Konvergenz ihrer Orthogonalentwicklungen, *Acta Sci. Math. Szeged,* **18** (1957), S. 131—139.

[5] S. Banach, Sur la divergence des séries orthogonales, *Studia Math.*, **9** (1940), S. 139—154.

[6] Н. К. Бари, О наилучщем приближении тригонометрическими полиномами двух сопряженных функций, Изв. Акад. Наук СССР, Сер. Мат., **19** (1955), S. 285—302.

[7] G. Freud, Über die starke $(C, 1)$-Summierbarkeit von orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 83—88.

[8] G. H. Hardy and J. E. Littlewood, Some properties of fractional integrals. I, *Math. Zeitschrift,* **27** (1928), S. 565—606.

[9] D. Králik, Untersuchung der Integrale und Derivierten gebrochener Ordnung mit den Methoden der konstruktiven Funktionentheorie, *Acta Math. Acad. Sci. Hung.*, **7** (1956), S. 49—64.

[10] D. Králik, Über die approximationstheoretische Charakterisierung gewisser Funktionenklassen, *Acta Math. Acad. Sci. Hung.*, **11** (1960), S. 377—386.

[11] J. Meder, On the estimation of Cesàro means of orthonormal series, *Annales Polonici Mathematici*, **4** (1957/58), S. 183—200.

[12] R. E. A. C. Paley and A. Zygmund, On some series of functions, *Proc. Cambridge Phil. Soc.*, **26** (1930), S. 337—357.

[13] K. Tandori, Über die orthogonalen Funktionen. VII (Approximationssätze), *Acta Sci. Math. Szeged,* **20** (1959), S. 19—24.

[14] A. Zygmund, *Trigonometric series. I—II* (Cambridge, 1959).

ON THE β-EXPANSIONS OF REAL NUMBERS

By

W. PARRY (London)

(*Presented by* A. RÉNYI)

Introduction

A. RÉNYI has proved [1]:

a) If $\beta > 1$, then every non-negative number x has a β-expansion:

$$x = \varepsilon_0(x) + \frac{\varepsilon_1(x)}{\beta} + \frac{\varepsilon_2(x)}{\beta^2} + \cdots \tag{1}$$

where $\varepsilon_0(x) = [x]$, $\varepsilon_1(x) = [\beta(x)]$, $\varepsilon_2(x) = [\beta(\beta(x))]$ etc. (Here and in what follows $[x]$ denotes the integral part and (x) the fractional part of the real number x.)

b) $T(x) = (\beta x)$ is an ergodic[1] transformation sending $[0, 1)$ onto itself, and for any $g(x) \in L[0, 1)$ we have for almost all x

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} g(T^k(x)) = M(g) \tag{2}$$

where $M(g)$ is a constant depending only on $g(x)$. Furthermore there exists a unique normalised measure ν equivalent to Lebesgue measure and invariant under T such that for any measurable subset E of $[0, 1)$

$$\nu E = \int_E h(x)\,dx \tag{3}$$

where $h(x)$ is a measurable function and

$$1 - \frac{1}{\beta} \leq h(x) \leq \frac{1}{1 - 1/\beta}$$

and

$$M(g) = \int_0^1 g(x) h(x)\,dx.$$

c) If β is the positive solution of $\beta^2 = \beta + 1$, then the unique normalised invariant measure corresponding to β may be represented as

$$\nu E = \int_E h(x)\,dx$$

1 RÉNYI's proof that T is ergodic, can easily be adopted to prove that $T \times T$ ([2], pp. 39—41) is ergodic and hence that T is weakly mixing.

where

$$h(x)=\begin{cases}\dfrac{5+3\sqrt{5}}{2} & \text{for } 0\leqq x<\dfrac{\sqrt{5}-1}{2},\\[2mm] \dfrac{5+\sqrt{5}}{2} & \text{for } \dfrac{\sqrt{5}-1}{2}\leqq x<1.\end{cases}$$

We call those β which have recurrent "tails" (i. e. $\varepsilon_{n+k}(\beta)=\varepsilon_n(\beta)$ for $n\geqq l$) in their β-expansions, β-numbers. Those with zero tails we call simple β-numbers. (RÉNYI's case c) is an example of a simple β-number.)

In § 1 we find, to within a constant factor of multiplication, for each $\beta>1$, an explicit determination of the unique measure invariant under the transformation $T(x)=(\beta x)$ and equivalent to Lebesgue measure. The derivative with respect to Lebesgue measure of this measure is a pure jump function ([3], p. 14), reducing to a step function with a finite number of steps precisely in the case of β-numbers.

The digits in β-expansions do not occur randomly (i. e. independently) and in § 2 we determine precisely the restrictions on the sequences of digits in β-expansions. We conclude this section by showing that the "conjugates" of each β-number with respect to its "characteristic equation" must all have absolute value less than 2.

Simple β-numbers are everywhere dense in $(1, \infty)$. This is proved in § 3.

Although we have an explicit formula for the invariant measure corresponding to each $\beta>1$, it does not in general yield normalised measures. Our attention in § 4 is therefore concentrated on the properties of the normalising factor as a function of β.

We show that this function is continuous from the right and the set of points of discontinuity from the left coincides with the set of simple β-numbers. Moreover, the normalising function tends to infinity as β tends to 1 and tends to 1 as β tends to infinity.

My thanks are due to YAEL NAIM DOWKER and H. HALBERSTAM for their many helpful suggestions.

§ 1. Invariant measures

For every $\beta>1$ there is a unique normalised measure ν equivalent to Lebesgue measure and invariant under $T(x)=(\beta x)$. By the Radon—Nikodym theorem [4], therefore, there is a measurable function $h_\nu(x)$, essentially unique, such that

$$\nu E=\int_E h_\nu(x)\,dx \tag{1.1}$$

for every Lebesgue set E. In the following $h_\nu(x)$ is defined to within a set of measure zero by (1.1). $x(m)$ denotes $\frac{x+m}{\beta}$.

THEOREM 1. *Let ν be a normalised measure equivalent to Lebesgue measure and let $T(x)=(\beta x)$ for $\beta>1$. A necessary and sufficient condition that $\nu E=\nu T^{-1}E$ for all Lebesgue sets E is that*

$$\beta h_\nu(x)=\sum_{Ty=x} h_\nu(y)=\sum_{m=0}^{[\beta-x]} h_\nu(x(m)) \quad p.p. \tag{1.2}$$

PROOF. Suppose $\nu E=\nu T^{-1}E$ for all Lebesgue sets E, then

$$\nu[a,b)=\int_a^b h_\nu(x)dx=\nu T^{-1}[a,b)=\sum_{m=0}^{[\beta-b]}\int_{a(m)}^{b(m)} h_\nu(x)dx$$

as long as $[\beta-b]$ is the largest integer m for which both $a(m)$ and $b(m)$ are less than 1, and this occurs when either $a<(\beta)$ and $b<(\beta)$, or $a>(\beta)$ and $b>(\beta)$. In this case, therefore,

$$\beta\left(\int_a^b h_\nu(x)dx/(b-a)\right)=\sum_{m=0}^{[\beta-b]}\left(\left(\int_{a(m)}^{b(m)} h_\nu(x)dx\right)\Big/(b(m)-a(m))\right).$$

We let a approach b and have for almost all b ([5], p. 284)

$$\beta h_\nu(b)=\sum_{m=0}^{[\beta-b]} h_\nu(b(m))=\sum_{Ty=b} h_\nu(y).$$

On the other hand, given for almost all x

$$\beta h_\nu(x)=\sum_{Ty=x} h_\nu(y) \tag{1.3}$$

we show that $\nu E=\nu T^{-1}E$ for all Lebesgue sets E.

We may suppose (1.3) holds everywhere by changing the values of $h_\nu(x)$ to zero where necessary.

If $[\beta-a]=[\beta-b]$, then

$$\int_a^b h_\nu(x)dx=\int_a^b\left(\frac{1}{\beta}\sum_{m=0}^{[\beta-x]} h_\nu(x(m))dx=\frac{1}{\beta}\sum_{m=0}^{[\beta-b]}\int_a^b h_\nu(x(m))dx=\right.$$

$$=\sum_{m=0}^{[\beta-b]}\int_{a(m)}^{b(m)} h_\nu(y)dy=\int_{T^{-1}[a,b]} h_\nu(y)dy.$$

Consequently $\nu[a, b) = \nu T^{-1}[a, b)$ for every interval $[a, b)$ with $[\beta - a] = [\beta - b]$, which happens when either a, b are both greater than (β) or both less than (β). This is sufficient by virtue of the countable additivity of T^{-1} over disjoint sets and the boundedness of $h_\nu(x)$ to establish the theorem.

For convenience we define $T^0(x) = x$ and inductively $T^n(1) = T^{n-1}((\beta))$ for $n \geqq 1$. Consequently $T^n(1) = (\beta T^{n-1}(1))$ for $n \geqq 1$, and

$$\beta = \varepsilon_0(\beta) + \frac{\varepsilon_1(\beta)}{\beta} + \frac{\varepsilon_2(\beta)}{\beta^2} + \cdots \tag{1.4}$$

where, by a),

$$\varepsilon_n(\beta) = [\beta T^{n-1}((\beta))] = [\beta T^n(1)].$$

Let us denote the pure jump function $\sum_{x < T^n(1)} \frac{1}{\beta^n}$ by $h_\beta(x)$ for each $\beta > 1$.

THEOREM 2. $h_\nu(x) = h_\beta(x)$ *p. p.*

PROOF. We show that $\beta h_\beta(x) = \sum_{m=0}^{[\beta - x]} h_\beta(x(m))$. Let

$$a_{n,m} = \begin{cases} 1 & \text{if} \quad x(m) < T^n(1), \\ 0 & \text{otherwise} \end{cases}$$

and

$$a_n = \begin{cases} 1 & \text{if} \quad x < T^n(1), \\ 0 & \text{otherwise.} \end{cases}$$

Evidently $\sum_{m=0}^{[\beta - x]} a_{n,m}$ is the number of m among $0, 1, \ldots, [\beta - x]$ with $x(m) < T^n(1)$. Therefore

$$\sum_{m=0}^{[\beta - x]} a_{n,m} = \begin{cases} [\beta T^n(1)] + 1 & \text{if} \quad x < (\beta T^n(1)) = T^{n+1}(1), \\ [\beta T^n(1)] & \text{otherwise.} \end{cases}$$

$$\sum_{m=0}^{[\beta - x]} h_\beta(x(m)) = \sum_{m=0}^{[\beta - x]} \sum_{n=0}^{\infty} \frac{a_{n,m}}{\beta^n} = \sum_{n=0}^{\infty} \sum_{m=0}^{[\beta - x]} \frac{a_{n,m}}{\beta^n} = \sum_{n=0}^{\infty} \frac{[\beta T^n(1)] + a_{n+1}}{\beta^n} =$$

$$= \sum_{n=0}^{\infty} \frac{\varepsilon_n(\beta)}{\beta^n} + \beta \sum_{n=0}^{\infty} \frac{a_{n+1}}{\beta^{n+1}} = \beta + \beta \sum_{n=0}^{\infty} \frac{a_n}{\beta^n} - \beta a_0 = \beta h_\beta(x).$$

COROLLARY. *$h_\beta(x)$ is a step function with a finite number of steps if and only if β has a recurrent tail in its β-expansion.*

PROOF. β has a recurrent tail in its β-expansion means, by a), $T^m(1) = T^n(1)$ for some $m \neq n$. Evidently the pure jump function $h_\beta(x)$ reduces to a step function if and only if $T^m(1) = T^n(1)$ for some $m \neq n$.

§ 2. β-numbers and their characteristic equations

If $(a_0, a_1, \ldots)$ and $(b_0, b_1, \ldots)$ are finite or infinite sequences of non-negative integers with the same number of terms, we write

$$(a_0, a_1, \ldots) < (b_0, b_1, \ldots)$$

when $a_n < b_n$ for the first n for which $a_n \neq b_n$. We also write $(a_0, a_1, \ldots)(\beta)$ for

$$a_0 + \frac{a_1}{\beta} + \frac{a_2}{\beta^2} + \cdots.$$

LEMMA 1. *If* $\beta = a_0 + \dfrac{a_1}{\beta} + \cdots$ *where* $(a_0, a_1, \ldots)$ *is a sequence of non-negative integers and if* $(b_0, b_1, \ldots)$ *is a sequence of non-negative integers with* $(b_n, b_{n+1}, \ldots) < (a_0, a_1, \ldots)$ *for all* $n \geqq 1$, *then*

$$\text{(2. 1)} \qquad \begin{cases} (b_n, b_{n+1}, \ldots)(\beta) < (a_m, a_{m+1}, \ldots)(\beta) \quad \textit{whenever} \\ (b_n, b_{n+1}, \ldots) < (a_m, a_{m+1}, \ldots), \end{cases}$$

unless $\beta = (a_0, a_1, \ldots, a_q)(\beta)$ *with some* q *and the tail of* $\{b_n\}$ *coincides with* $\{c_n\}$.

$$\left(c_i = \begin{cases} a_i & \textit{if} \quad i \not\equiv 0 \mod q, \\ a_q - 1 & \textit{if} \quad i \equiv 0 \mod q. \end{cases} \right)$$

PROOF. We first prove that $(b_n, b_{n+1}, \ldots)(\beta) \leqq (a_m, a_{m+1}, \ldots)(\beta)$ whenever $(b_n, b_{n+1}, \ldots) < (a_m, a_{m+1}, \ldots)$ by showing

$$\text{(2. 2)} \qquad \begin{cases} (b_n, \ldots, b_{n+r})(\beta) < (a_m, \ldots, a_{m+r})(\beta) \quad \text{whenever} \\ (b_n, \ldots, b_{n+r}) < (a_m, \ldots, a_{m+r}). \end{cases}$$

In the case $r = 0$, $(b_n) < (a_m)$ implies $(b_n)(\beta) = (b_n) < (a_m) = (a_m)(\beta)$. Therefore (2. 2) holds for all m, n where $r = 0$. We suppose (2. 2) holds for all m, n when $r < k$.

If $(b_n, \ldots, b_{n+k}) < (a_m, \ldots, a_{m+k})$, either $b_n = a_m$ and $(b_{n+1}, \ldots, b_{n+k}) < < (a_{m+1}, \ldots, a_{m+k})$ (in which case $(b_{n+1}, \ldots, b_{n+k})(\beta) < (a_{m+1}, \ldots, a_{m+k})(\beta)$ by hypothesis) or $b_n < a_m$. In the latter case, since $(b_{n+1}, \ldots, b_{n+k}) \leqq (a_0, \ldots, a_{k-1})$ and again by hypothesis, we have

$$\text{(2. 3)} \qquad \begin{aligned} &(b_n, \ldots, b_{n+k})(\beta) \leqq (b_n, a_0, \ldots, a_{k-1})(\beta) \leqq \\ &\leqq (a_m, 0, \ldots, 0)(\beta) \leqq (a_m, \ldots, a_{m+k})(\beta). \end{aligned}$$

The central inequality in (2. 3) holds by virtue of the relation

$$\frac{a_0}{\beta} + \cdots + \frac{a_{k-1}}{\beta^k} \leqq a_m - b_n.$$

We note that equality holds throughout (2. 3) when and only when

$$\beta=a_0+\frac{a_1}{\beta}+\cdots+\frac{a_{k-1}}{\beta^{k-1}} \quad\text{and}\quad (b_{n+1},\ldots,b_{n+k})=(a_0,\ldots,a_{k-1}).$$

However, this is excluded by the condition $(b_n, b_{n+1},\ldots)<(a_0,a_1,\ldots)$ for all $n\geqq 1$.

Consequently (2. 2) is proved and it follows immediately that

$$(b_n, b_{n+1},\ldots)(\beta)\leqq(a_m, a_{m+1},\ldots)(\beta)$$

whenever

$$(b_n, b_{n+1},\ldots)<(a_m, a_{m+1},\ldots).$$

Suppose $(b_n, b_{n+1},\ldots)<(a_m, a_{m+1},\ldots)$, then for some p, q, $p-n=q-m\geqq 0$

$$(b_p, b_{p+1},\ldots)(\beta)\leqq(b_p, a_0, a_1,\ldots)(\beta)\leqq(a_q, 0,\ldots,0)(\beta)\leqq(a_q, a_{q+1},\ldots)(\beta)$$

and strict equality holds throughout only if

$$a_{q+1}=a_{q+2}=\cdots=0.$$

In this case $\beta=(a_0, a_1,\ldots,a_q)(\beta)=(c_0, c_1,\ldots)(\beta)$ where

$$c_i=\begin{cases} a_i & \text{if } i\not\equiv 0 \mod q,\\ a_q-1 & \text{if } i\equiv 0 \mod q,\end{cases}$$

and consequently $(b_n, b_{n+1},\ldots)(\beta)<(c_m, c_{m+1},\ldots)(\beta)=(a_m,\ldots)(\beta)$ whenever $(b_n, b_{n+1},\ldots)<(c_m, c_{m+1},\ldots)$. But $(b_n, b_{n+1},\ldots)<(a_m, a_{m+1},\ldots)$ implies either the tail of $\{b_n\}$ coincides with $\{c_n\}$, or $(b_n, b_{n+1},\ldots)<(c_m, c_{m+1},\ldots)$.

This completes the proof.

THEOREM 3. *If $\beta>1$ and the β-expansion of β is*

$$\beta=a_0+\frac{a_1}{\beta}+\frac{a_2}{\beta^2}+\cdots$$

and if $(b_0, b_1,\ldots)$ is a sequence of non-negative numbers whose tail, in case the β-expansion of β is finite, does not coincide with $\{c_n\}$ (cf. Lemma 1), *a necessary and sufficient condition for the existence of x with β-expansion*

$$x=b_0+\frac{b_1}{\beta}+\cdots$$

is that $(b_n, b_{n+1},\ldots)<(a_0, a_1,\ldots)$ for all $n\geqq 1$.

In particular, $(a_n, a_{n+1},\ldots)<(a_0, a_1,\ldots)$ for all $n\geqq 1$.

PROOF. If there exists x with β-expansion

$$x=b_0+\frac{b_1}{\beta}+\cdots$$

for every $n \geqq 1$,

$$\frac{b_n}{\beta} + \frac{b_{n+1}}{\beta^2} + \cdots < 1 = \frac{a_0}{\beta} + \frac{a_1}{\beta^2} + \cdots.$$

Consequently for every $n \geqq 1$

$$(b_n, b_{n+1}, \ldots) \neq (a_0, a_1, \ldots).$$

Let k be the first integer with $b_{n+k} \neq a_k$, then

$$b_{n+k} + \frac{b_{n+k+1}}{\beta} + \cdots < a_k + \frac{a_{k+1}}{\beta} + \cdots,$$

and since $\frac{a_{k+1}}{\beta} + \frac{a_{k+2}}{\beta^2} + \cdots < 1$, we have $b_{n+k} < a_k + 1$, i. e. $b_{n+k} < a_k$.

By Lemma 1, for all $n \geqq 1$, $(b_n, b_{n+1}, \ldots) < (a_0, a_1, \ldots)$ implies

$$\frac{b_n}{\beta} + \frac{b_{n+1}}{\beta^2} + \cdots < \frac{a_0}{\beta} + \frac{a_1}{\beta^2} + \cdots = 1,$$

and there exists x with β-expansion

$$x = b_0 + \frac{b_1}{\beta} + \frac{b_2}{\beta^2} + \cdots.$$

In the following corollaries $(a_0, a_1, \ldots)$ is a sequence of non-negative integers where $a_0 \geqq 1$ and $a_n \leqq a_0$ for all $n \geqq 0$.

It should be noted that under this condition there is a unique solution $\beta > 1$ of the equation

$$x = a_0 + \frac{a_1}{x} + \cdots$$

and this solution lies between a_0 and $a_0 + 1$.

By Lemma 1 and Theorem 3, we have

COROLLARY 1. *The unique solution $\beta > 1$ of*

$$x = a_0 + \frac{a_1}{x} + \cdots$$

has for its β-expansion

$$\beta = a_0 + \frac{a_1}{\beta} + \cdots$$

if and only if $(a_n, a_{n+1}, \ldots) < (a_0, a_1, \ldots)$ for all $n \geqq 1$.

We omit the proofs of the following corollaries:

COROLLARY 2. *The unique solution* $\beta>1$ *of*

$$x=\sum_{k=0}^{\infty}\frac{a_k}{x^k} \tag{2.4}$$

(where $a_{n+rm+k}=a_{n+k}$, $r\geqq 0$, $0\leqq k<m$*) has*

$$\beta=\sum_{k=0}^{\infty}\frac{a_k}{\beta^k}$$

for its β*-expansion if and only if*

$$\begin{cases}(a_0,\dots,a_{n+m-1})>(a_1,\dots,a_{n+m-1},a_{n+1}),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ (a_0,\dots,a_{n+m-1})>(a_k,\dots,a_{n+r(k)}),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ (a_0,\dots,a_{n+m-1})>(a_{n+m-1},a_{n+r(n+m-1)})\\ \qquad \text{where} \quad r(k)\equiv k \mod m.\end{cases} \tag{2.5}$$

COROLLARY 3. *The unique solution* $\beta>1$ *of*

$$x^{n+m}-(a_0x^{n+m-1}+\cdots+a_{n+m-1})=x^n-(a_0x^{n-1}+\cdots+a_{n-1})$$

(where $a_{n+rm+k}=a_{n+k}$, $r\geqq 0$, $0\leqq k<m$*) has*

$$\beta=\sum_{k=0}^{\infty}\frac{a_k}{\beta^k}$$

for its β*-expansion if and only if* (2.5) *holds.*

COROLLARY 4. *The unique solution* $\beta>1$ *of*

$$x=a_0+\frac{a_1}{x}+\cdots+\frac{a_n}{x^n}$$

has for its β*-expansion*

$$\beta=a_0+\frac{a_1}{\beta}+\cdots+\frac{a_n}{\beta^n}$$

if and only if

$$\begin{cases}(a_0,\dots,a_n)>(a_1,\dots,a_n,0),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ (a_0,\dots,a_n)>(a_k,\dots,a_n,0,\dots,0),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ (a_0,\dots,a_n)>(a_n,0,\dots,0).\end{cases} \tag{2.6}$$

(2.6) *is equivalent to*

$$\begin{cases}(a_0,\dots,a_{n-1})\geqq(a_1,\dots,a_n),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ (a_0,\dots,a_{n-k})\geqq(a_k,\dots,a_n),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \qquad (a_0)\geqq(a_n).\end{cases} \tag{2.7}$$

If β is a non-simple β-number with β-expansion

$$\beta=\sum_{k=0}^{\infty}\frac{a_k}{\beta^k}$$

(where $a_{n+rm+k}=a_{n+k}$, $r\geqq 0$, $0\leqq k<m$), the equation

$$z^{n+m}-(a_0z^{n+m-1}+\cdots+a_{n+m-1})=z^n-(a_0z^{n-1}+\cdots+a_{n-1})$$

is termed the characteristic equation of β.

Likewise if β is a simple β-number with β-expansion

$$\beta=a_0+\frac{a_1}{\beta}+\cdots+\frac{a_n}{\beta^n},$$

the equation

$$z^{n+1}-(a_0z^n+\cdots+a_n)=0$$

is termed the characteristic equation of β.

LEMMA 2. *If β is a simple (non-simple) β-number with characteristic equation*

$$z^{n+m}-(\varepsilon_0z^{n+m-1}+\cdots+\varepsilon_{n+m-1})=0$$
$$(=z^n-(\varepsilon_0z^{n-1}+\cdots+\varepsilon_{n-1})),$$

the conjugates $\beta_1,\ldots,\beta_{n+m-1}$ of β with respect to the characteristic equation, satisfy

$$z^{n+m-1}+T(1)z^{n+m-2}+\cdots+T^{n+m-1}(1)=0 \tag{2.8}$$
$$(=z^{n-1}+T(1)z^{n-2}+\cdots+T^{n-1}(1)).$$

PROOF. This is easily verified by multiplying both sides of (2.8) by $(z-\beta)$ and adding $T^{n+m}(1)=0(=T^n(1))$ to both sides. This gives the characteristic equation of β.

THEOREM 4. *The conjugates of a β-number with respect to its characteristic equation have absolute value less than 2.*[2]

PROOF. Using the notation employed in Lemma 2 the conjugates of β with respect to its characteristic equation satisfy

$$z^{n+m-1}+T(1)z^{n+m-2}+\cdots+T^{n+m-1}(1)=0$$
$$(=z^{n-1}+\cdots+T^{n-1}(1)).$$

If $|z|>1$,

$$(|z|^m-1)|z^{n-1}+T(1)z^{n-2}+\cdots+T^{n-1}(1)|<|z|^{m-1}+\cdots+1,$$

[2] In this respect β-numbers resemble Pisot—Vijayaraghavan numbers ([5], pp. 133—144) defined as those algebraic integers greater than 1, with conjugates of maximum absolute value less than one.

i. e.

$$|z^{n-1}+T(1)z^{n-2}+\cdots+T^{n-1}(1)|<\frac{1}{|z|-1}.$$

Therefore

$$|z|^{n-1}<\frac{1}{|z|-1}+|T(1)z^{n-2}+\cdots+T^{n-1}(1)|\leqq$$

$$\leqq\frac{1}{|z|-1}+\frac{|z|^{n-1}-1}{|z|-1}=\frac{|z|^{n-1}}{|z|-1},$$

i. e.

$$|z|<2.$$

§ 3. Distribution of β-numbers

LEMMA 3. *If the α-expansion of α is*

$$\alpha=a_0+\frac{a_1}{\alpha}+\cdots$$

and the β-expansion of β is

$$\beta=b_0+\frac{b_1}{\beta}+\cdots,$$

then $\alpha>\beta$ if and only if $(a_0, a_1, \ldots)>(b_0, b_1, \ldots)$.

PROOF. Suppose on the contrary $\alpha<\beta$ and

$$(a_0, a_1, \ldots)>(b_0, b_1, \ldots).$$

(The case of equality is trivially dismissed.)

If n is the first integer with $a_n>b_n$,

$$b_n+\frac{b_{n+1}}{\beta}+\cdots<b_n+1\leqq a_n<a_n+\frac{a_{n+1}}{\beta}+\cdots,$$

therefore

$$\beta=b_0+\frac{b_1}{\beta}+\cdots<a_0+\frac{a_1}{\beta}+\cdots<a_0+\frac{a_1}{\alpha}+\frac{a_2}{\alpha^2}+\cdots=\alpha.$$

This contradiction completes the proof.

THEOREM 5. *The set of simple β-numbers is everywhere dense in* $(1, \infty)$.

PROOF. Every $\beta>1$ has a β-expansion

$$\beta=\varepsilon_0+\frac{\varepsilon_1}{\beta}+\cdots$$

where

$$\varepsilon_n = \varepsilon_n(\beta) \qquad (n = 0, 1, \dots)$$

and according to Lemma 2

$$(\varepsilon_0, \varepsilon_1, \dots) > (\varepsilon_n, \varepsilon_{n+1}, \dots) \quad \text{for} \quad n \geqq 1$$

and therefore

$$(\varepsilon_0, \dots, \varepsilon_n) > (\varepsilon_1, \dots, \varepsilon_n, 0),$$
$$\dots\dots\dots\dots\dots\dots$$
$$(\varepsilon_0, \dots, \varepsilon_n) > (\varepsilon_n, 0, \dots, 0).$$

Consequently, by Corollary 4 of Theorem 3, there is a simple β-number $\beta(n)$ with $\beta(n)$-expansion

$$\beta(n) = (\varepsilon_0, \varepsilon_1, \dots, \varepsilon_n)\,(\beta(n)).$$

It suffices therefore to show that $\lim\limits_{n\to\infty} \beta(n) = \beta \cdot \beta(n)$ is non-decreasing as n increases (Lemma 3).

Suppose there exists K with $\beta(n) < K < \beta$ for all n, then $\beta(n) > \varepsilon_0 + + \frac{\varepsilon_1}{K} + \cdots + \frac{\varepsilon_n}{K^n}$ for all n and therefore

$$K \geqq \lim_{n\to\infty} \beta(n) \geqq \varepsilon_0 + \frac{\varepsilon_1}{K} + \cdots > \varepsilon_0 + \frac{\varepsilon_1}{\beta} + \cdots = \beta.$$

This contradiction shows that $\lim\limits_{n\to\infty} \beta(n) = \beta$ and this completes the proof.

§ 4. The normalising function

The most important reason for evaluating the unique normalised invariant measure equivalent to Lebesgue measure is to find for each β the frequency of the digits $0, 1, \dots, [\beta]$ in the β-expansion of each x.

It can be shown using BIRKHOFF's ergodic theorem [7] that the frequency of each digit exists and is the same for almost all x and is expressed by the function $\Phi(n)$:

$$\Phi(n) = \begin{cases} \nu\left[\frac{n}{\beta}, \frac{n+1}{\beta}\right) & \text{if} \quad n < [\beta], \\ \nu\left[\frac{[\beta]}{\beta}, 1\right) & \text{if} \quad n = [\beta] \end{cases} \tag{4.1}$$

where $\nu E = \frac{1}{F(\beta)} \int\limits_E h_\beta(x)\,dx$ and $F(\beta)$ is the normalising factor

$$\int\limits_0^1 h_\beta(x)\,dx.$$

We note the following formulae:

$$F(\beta)=\int_0^1 h_\beta(x)\,dx\int_0^1\Big(\sum_{x<T^n(1)}\frac{1}{\beta^n}\Big)dx=\int_0^1\Big(\sum_{n=0}^{\infty}\frac{a_n(x)}{\beta^n}\Big)dx,$$

where

$$a_n(x)=\begin{cases}1 & \text{if } x<T^n(1),\\ 0 & \text{otherwise,}\end{cases}$$

$$(4.2)\qquad F(\beta)=\sum_{n=0}^{\infty}\frac{1}{\beta^n}\int_0^1 a_n(x)\,dx=\sum_{n=0}^{\infty}\frac{T^n(1)}{\beta^n}=\sum_{n=0}^{\infty}\frac{(n+1)\varepsilon_n(\beta)}{\beta^n}$$

since $T^n(1)=\frac{\varepsilon_n(\beta)}{\beta}+\frac{\varepsilon_{n+1}(\beta)}{\beta^2}+\cdots$ for all $n\geqq 0$.

We have therefore $F(n)=1$ for each integer $n>1$, and

$$1\leqq F(\beta)\leqq 1+\frac{1}{\beta}+\frac{1}{\beta^2}+\cdots=\frac{\beta}{\beta-1},$$

and consequently

$$(4.3)\qquad \lim_{\beta\to\infty}F(\beta)=1.$$

LEMMA 4. *If the β-expansion of β is*

$$\beta=\varepsilon_0+\frac{\varepsilon_1}{\beta}+\cdots+\frac{\varepsilon_{n-1}}{\beta^{n-1}},$$

for every r there exists $\alpha=\alpha(r)$ with α-expansion

$$(4.4)\qquad \alpha(r)=(a_0,\ldots,a_{(r+1)n-1})(\alpha(r))$$

where

$$\begin{cases}a_{sn+t}=\varepsilon_t & \textit{for}\quad 0\leqq s\leqq r,\ 0\leqq t<n-1,\\ a_{sn+n-1}=\varepsilon_{n-1}-1 & \textit{for}\quad 0\leqq s\leqq r,\end{cases}$$

and $\alpha(r)\uparrow\beta$ as $r\to\infty$.

PROOF.

$$(\varepsilon_0,\ldots,\varepsilon_{n-1})>(\varepsilon_k,\ldots,\varepsilon_{n-1},0,\ldots,0),$$

i. e.

$$(\varepsilon_0,\ldots,\varepsilon_{n-k-1})>(\varepsilon_k,\ldots,\varepsilon_{n-1})>(\varepsilon_k,\ldots,\varepsilon_{n-1}-1).$$

Therefore

$$(a_0,\ldots,a_{(r+1)n-1})>(a_k,\ldots,a_{(r+1)n-1},0,\ldots,0)$$

where the a_i's are defined in (4.4).

By Corollary 4 of Theorem 3 there exists $\alpha=\alpha(r)$ with α-expansion (4. 4). Evidently

$$\alpha=(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}-1)(\alpha)\left(1+\frac{1}{\alpha^n}+\cdots+\frac{1}{\alpha^{rn}}\right)=$$

$$=(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}-1)(\alpha)\left(\frac{1-1/\alpha^{(r+1)n}}{1-1/\alpha^n}\right).$$

According to Lemma 3 $\alpha(r)$ is non-decreasing as r increases. Let $\gamma=\sup\limits_r \alpha(r)$, then

$$\gamma=(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}-1)(\gamma)\left(\frac{1}{1-1/\gamma^n}\right),$$

i. e.

$$\gamma=(\varepsilon_0, \ldots, \varepsilon_{n-1})(\gamma),$$

and therefore $\gamma=\beta$.

THEOREM 6. *If β is a simple β-number with β-expansion*

$$\beta=\varepsilon_0+\frac{\varepsilon_1}{\beta}+\cdots+\frac{\varepsilon_{n-1}}{\beta^{n-1}},$$

then

$$\lim_{\gamma\uparrow\beta} F(\gamma)=F(\beta)\left(\frac{\beta^n}{\beta^n-1}\right).$$

PROOF. If $\alpha=\alpha(r)\leqq\gamma<\beta$ ($\alpha(r)$ defined in (4. 4)), γ will have in its γ-expansion the same $(r+1)n$ initial digits as $\alpha(s)$ for all $s\geqq r$. (By Lemma 3.)

Therefore

$$\gamma=(\varepsilon_0, \ldots, \varepsilon_{n-1}-1)(\gamma)+\frac{1}{\gamma^n}(\varepsilon_0, \ldots, \varepsilon_{n-1}-1)(\gamma)+\cdots$$

$$\cdots+\frac{1}{\gamma^{rn}}(\varepsilon_0, \ldots, \varepsilon_{n-1}-1)(\gamma)+\cdots$$

and

$$G_r(\beta)\equiv\frac{1}{\beta}(\varepsilon_0, 2\varepsilon_1, \ldots, n(\varepsilon_{n-1}-1))(\beta)+$$

$$+\frac{1}{\beta^{n+1}}((n+1)\varepsilon_0, (n+2)\varepsilon_1, \ldots, 2n(\varepsilon_{n-1}-1))(\beta)+\cdots$$

$$\cdots+\frac{1}{\beta^{rn+1}}((rn+1)\varepsilon_0, \ldots, (r+1)n(\varepsilon_{n-1}-1))(\beta)\leqq F(\gamma)\leqq$$

$$\leqq G_r(\gamma)+\theta_r\leqq G_r(\alpha(r))+\theta_r$$

where $\lim\limits_{r\to\infty}\theta_r=0$.

If suffices to show:[8]

$$\lim_{r\to\infty} F(\alpha(r)) = \lim_{r\to\infty} G_r(\alpha) = F(\beta)\left(\frac{\beta^n}{\beta^n-1}\right).$$

Since

$$G_r(\alpha)-\left(\frac{\varepsilon_0}{\alpha}+\frac{2\varepsilon_1}{\alpha^2}+\cdots+\frac{n(\varepsilon_{n-1}-1)}{\alpha^n}\right)=\frac{1}{\alpha^{n+1}}((n+1)\varepsilon_0,\ldots,2n(\varepsilon_{n-1}-1))(\alpha)+$$

$$+\cdots+\frac{1}{\alpha^{rn+1}}((rn+1)\varepsilon_0,\ldots,(r+1)n(\varepsilon_{n-1}-1))(\alpha)=$$

$$=\frac{1}{\alpha^n}\left[n\left\{\frac{1}{\alpha}(\varepsilon_0,\ldots,\varepsilon_{n-1}-1)(\alpha)+\cdots\right.\right.$$

$$\left.\left.\cdots+\frac{1}{\alpha^{(r-1)n+1}}(\varepsilon_0,\ldots,\varepsilon_{n-1}-1)(\alpha)\right\}+G_{r-1}(\alpha)\right]=\frac{1}{\alpha^n}[n\,T^n(1)+G_{r-1}(\alpha)],$$

we have

$$\lim_{r\to\infty} T_r^n(1)=\lim_{r\to\infty}\left\{\frac{1}{\alpha}\left(\varepsilon_0+\frac{\varepsilon_1}{\alpha}+\cdots+\frac{\varepsilon_{n-1}-1}{\alpha^{n-1}}\right)\left(\frac{1-1/\alpha^{rn}}{1-1/\alpha^n}\right)\right\}=$$

$$=\frac{1}{\beta}\left(\varepsilon_0+\frac{\varepsilon_1}{\beta}+\cdots+\frac{\varepsilon_{n-1}-1}{\beta^{n-1}}\right)\left(\frac{1}{1-1/\beta^n}\right)=1$$

and

$$\lim_{r\to\infty}\left(\frac{\varepsilon_0}{\alpha}+\frac{2\varepsilon_1}{\alpha^2}+\cdots+\frac{n(\varepsilon_{n-1}-1)}{\alpha^n}\right)=\frac{\varepsilon_0}{\beta}+\frac{2\varepsilon_1}{\beta^2}+\cdots+\frac{n(\varepsilon_{n-1}-1)}{\beta^n},$$

therefore

$$\lim_{r\to\infty} G_r(\alpha)\left(1-\frac{1}{\alpha^n}\right)=\frac{\varepsilon_0}{\beta}+\cdots+\frac{n\varepsilon_{n-1}}{\beta^n},$$

i. e.

$$\lim_{r\to\infty} G_r(\alpha)=F(\beta)\left(\frac{\beta^n}{\beta^n-1}\right).$$

THEOREM 7. *If* $\beta>1$, *then* $\lim_{\gamma\downarrow\beta} F(\gamma)=F(\beta)$, *and if* β *is not a simple* β*-number,* $\lim_{\gamma\uparrow\beta} F(\gamma)=F(\beta)$.

PROOF. In either case suppose that $|\gamma-\beta|<\delta$ and let $n=n(\delta)$ be the number of initial consecutive digits common to both the γ-expansion of γ

[8] In what follows, the dependence of α on r should be noted.

and the β-expansion of β. By Lemma 4 it is easy to see that $\lim_{\delta\to\infty} n(\delta)=\infty$,

$$F(\beta)=\frac{\varepsilon_0}{\beta}+\frac{2\varepsilon_1}{\beta^2}+\cdots+\frac{n\varepsilon_{n-1}}{\beta^n}+\theta_n \quad \text{where} \quad \lim_{n\to\infty}\theta_n=0,$$

$$F(\gamma)=\frac{\varepsilon_0}{\gamma}+\frac{2\varepsilon_1}{\gamma^2}+\cdots+\frac{n\varepsilon_{n-1}}{\gamma^n}+\psi_n \quad \text{where} \quad \lim_{n\to\infty}\psi_n=0,$$

$$|F(\beta)-F(\gamma)|\leqq\left|\frac{1}{\beta}-\frac{1}{\gamma}\right|\left(\varepsilon_0+2\varepsilon_1\left(\frac{1}{\beta}+\frac{1}{\gamma}\right)+\cdots+n\varepsilon_{n-1}\left(\frac{1}{\beta^{n-1}}+\cdots+\frac{1}{\gamma^{n-1}}\right)\right)+$$

$$+\theta_n+\psi_n\leqq\left|\frac{1}{\beta}-\frac{1}{\gamma}\right|\left(\varepsilon_0+\frac{2^2\varepsilon_0}{\alpha}+\cdots+\frac{n^2\varepsilon_0}{\alpha^{n-1}}\right)+\theta_n+\psi_n \quad \text{where} \quad \alpha=\min(\beta,\gamma).$$

Therefore $\lim_{\gamma\uparrow\beta}F(\gamma)=F(\beta)$ if β is not a simple β-number and $\lim_{\gamma\downarrow\beta}F(\gamma)=F(\beta)$ for all $\beta>1$.

THEOREM 8. $\lim_{\gamma\downarrow 1}F(\gamma)=\infty$.

PROOF. Let us write $\beta=\beta(n)$ for the positive solution of

$$\beta=1+\frac{1}{\beta^n}.$$

$\beta(n)$ decreases as n increases (Lemma 4). Should there exist K with $\beta(n)\geqq K>1$ for all n, we would have

$$\lim_{n\to\infty}\beta(n)\leqq\lim_{n\to\infty}\left(1+\frac{1}{K^n}\right)=1.$$

Consequently $\lim_{n\to\infty}\beta(n)=1$.

Therefore, by virtue of the continuity from the right of $F(\beta)$ at points $\beta>1$, it suffices to show that $\lim_{n\to\infty}F(\beta(n))=\infty$.

$$F(\beta(n))=\frac{1}{\beta(n)}+\frac{n+1}{\beta(n)^n}$$

and we show that

$$\lim_{n\to\infty}\frac{n}{\beta(n)^n}=\infty.$$

Suppose $\frac{n}{\beta(n)^n}<K$ for infinitely many n, then

$$\beta(n)^n=\left(1+\frac{1}{\beta(n)^n}\right)^n<\left(1+\frac{1}{\beta(n)^n}\right)^{K\beta(n)^n}$$

for infinitely many n.

But $\left(1+\frac{1}{\beta(n)^n}\right)^{K\beta(n)^n}$ converges to e^K, since $\lim_{n\to\infty}\beta(n)^n=\infty$ by virtue of the relation

$$\beta(n)^n=\frac{1}{\beta(n)-1}.$$

Therefore there exists arbitrarily large n with $\beta(n)^n<e^K$ and yet $\lim_{n\to\infty}\beta(n)^n=\infty$.

This contradiction establishes the theorem.

(Received 24 February 1960)

References

[1] A. Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hung.*, **8** (1957), pp. 477—493.
[2] P. R. Halmos, *Lectures on ergodic theory*, Mathematical Society of Japan (1956).
[3] F. Riesz et B. Sz.-Nagy, *Leçons d'analyse fonctionnelle* (Budapest, 1952).
[4] P. R. Halmos, *Measure theory* (New York, 1950).
[5] M. E. Munroe, *Introduction to measure and integration* (Cambridge, 1953).
[6] J. W. S. Cassels, *An introduction to diophantine approximations*, Cambridge Tracts (1957).
[7] G. D. Birkhoff, Proof of the ergodic theorem, *Proc. Nat. Acad. Sci. USA*, **17** (1931), pp. 162—166.

NOTE ON MIXING SEQUENCES OF EVENTS

By

L. SUCHESTON (Rochester, N. Y., USA)

(*Presented by* A. RÉNYI)

1. Introduction. Let Ω be a set, $\mathcal{A}$ a σ-field of subsets of Ω (called events), $\mathbf{P}$ a probability measure defined on $\mathcal{A}$.

DEFINITION 1. A sequence of events A_n $(n=1, 2, \ldots)$ is called *mixing with density* p $(0 \leqq p \leqq 1)$, if for each non-null event M

$$\mathbf{P}(A_n/M) = \mathbf{P}(A_n M)/\mathbf{P}(M) \to p. \tag{1}$$

Mixing sequences were studied in [1] and, independently, in [2][1] and [3] where they were called *equidistributed.* (We prefer the term mixing for its brevity and for its ergodic connotations.) We show here that RÉNYI's Theorem 2 [1], obtained by a Hilbert space argument, is contained in a special case of a result of [2], proven directly; also a Hilbert space proof of this special case is given. We then consider the following generalization of the notion of mixing:

DEFINITION 2 ([2]). A sequence of events A_n is called *equidistributed of order* r and *with density* p or *equidistributed* (r, p), if for each non-null event M, each subsequence of A_n contains a further subsequence B_n with

$$\lim_{\substack{n_1 \to \infty \\ n_1 < \cdots < n_r}} \mathbf{P}(B_{n_1} \ldots B_{n_r}/M) = p^r. \tag{2}$$

Another rather natural generalization of mixing is the following notion:

DEFINITION 3. A sequence of events A_n is called *mixing of order* r and *with density* p, or *mixing* (r, p), if for each non-null event M

$$\lim_{\substack{n_1 \to \infty \\ n_1 < \cdots < n_r}} \mathbf{P}(A_{n_1} \ldots A_{n_r}/M) = p^r. \tag{3}$$

For $r=1$ Definitions 2 and 3 clearly coalesce with Definition 1. Improving Theorem B [2] we show here that equidistribution (r, p) coincides with

[1] We take this opportunity to make the following corrections in [2]: P. 389, line 2: read Section 2 instead of [2]. P. 391, line 15: replace $q < m^{1/r}p$ by $q \leqq p$; line 16: delete $m^{1/r}$; line 17: replace p^s by $m^{-1}p^s$; line 8 from the bottom: delete "$0 \leqq x < 1$, and".

mixing with density p for *all* r. A different situation prevails for *mixing* of order r: If a sequence is mixing (r, p), then it is mixing (s, p) for all $s < r$, but not necessarily for an $s > r$.

2. A Hilbert space lemma. The restriction of Lemma 4 [2] to the case $r = 1, s = 2$ yields a result called here for the sake of reference:

THEOREM 1. *If A_n is a sequence of events such that*

$$\liminf \mathbf{P}(A_n) \geqq p, \tag{4}$$

then either the sequence A_n is mixing with density p or there exists a subsequence B_n of A_n such that for some $\delta > 0$

$$\mathbf{P}(B_i B_j) \geqq p^2 + \delta \qquad (i, j = 1, 2, \ldots). \tag{5}$$

It is easy to see that Theorem 1 implies the following Theorem 2 [1]:

THEOREM 2 (RÉNYI). *A sequence of events A_n such that*

$$\lim \mathbf{P}(A_n) = \lim \mathbf{P}(A_n/A_k) = p \qquad (k = 1, 2, \ldots) \tag{6}$$

is mixing with density p.

Let, indeed, A_n be a sequence satisfying (6). If B_n is an arbitrary subsequence of A_n, then $\mathbf{P}(B_i B_j) = \mathbf{P}(B_i)\mathbf{P}(B_j/B_i)$ can be made arbitrarily close to p^2, if at first i is chosen large, then for i fixed, j is chosen large. Therefore the sequence B_n cannot satisfy (5); it follows that Theorem 1 implies Theorem 2.

Theorem 2 is proven in [1] by application of a Hilbert space lemma:

LEMMA 1 (RÉNYI). *If f_n is a bounded sequence of elements of a Hilbert space and if*

$$\lim_{n \to \infty} (f_n, f_k) = 0 \qquad (k = 1, 2, \ldots), \tag{7}$$

then the sequence f_n converges weakly to zero.

We give here a Hilbert space lemma which allows a derivation of Theorem 1 above similar to the derivation in [1] of Theorem 2 from Lemma 1.

LEMMA 2. *If f_n is a bounded sequence of elements of a Hilbert space, then either the sequence f_n converges weakly to zero or there exists a subsequence g_n of f_n such that for some $\delta > 0$*

$$|(g_i, g_j)| \geqq \delta \qquad (i, j = 1, 2, \ldots). \tag{8}$$

PROOF. It is easy to see that if Lemma 2 holds with the additional condition $i \neq j$ added in (8) — (8) thus modified will be referred to as (8′) —, then this lemma holds in the form in which it is stated. If, indeed,

f_n does not converge weakly to zero, it is possible to consider a subsequence of elements of f_n whose norms admit a positive lower bound, and to apply to this subsequence Lemma 2 with (8) replaced by (8′). We shall show that if (8′) fails for every subsequence g_n of f_n, then every weakly converging subsequence of f_n converges weakly to zero. By the known weak compactness of a sphere in a Hilbert space (see f. i. [4], p. 97), every subsequence of f_n contains a weakly converging subsequence; it will follow that the sequence f_n itself converges weakly to zero.

Let h_n be a subsequence of f_n converging weakly to some element f of the Hilbert space. If (8′) fails for every subsequence g_n of f_n, then for every $\delta > 0$ every subsequence l_n of h_n admits at least two terms, say l_i, l_j such that

(9) $$|(l_i, l_j)| < \delta \qquad (i \neq j).$$

It follows by RAMSEY's theorem [5] that there is a subsequence m_n of h_n such that

(10) $$|(m_i, m_j)| < \delta \qquad (i, j = 1, 2, \dots; i \neq j).$$

Consider now a sequence of numbers δ_n decreasing to zero and apply the diagonal construction. One obtains a subsequence o_i of h_i such that

(11) $$\lim_{i,j\to\infty} (o_i, o_j) = 0 \qquad (i \neq j).$$

On the other hand, the expression

$$|(f, f) - (o_i, o_j)| \leqq |(f, f) - (f, o_j)| + |(f, o_j) - (o_i, o_j)|$$

can be made arbitrarily small if at first j is chosen large, then for j fixed, i is chosen large. It follows that $\|f\| = 0$. This proves Lemma 2.

3. Mixing sequences. For mixing sequences the following theorem holds:

THEOREM 3 (LORENTZ).[2] *If a sequence of events A_n is mixing with density p, then for each non-null event M and each $\varepsilon > 0$ there exists a subsequence B_n of A_n such that*

(12) $$(1-\varepsilon)p^r < \mathbf{P}(B_{n_1} \dots B_{n_r}/M) < (1+\varepsilon)p^r$$
$$(n_1 < \dots < n_r;\ n_1, \dots, n_r = 1, 2, \dots;\ r = 1, 2, \dots).$$

[2] Contained in a result stated in [3], p. 208. LORENTZ's paper is written in terms of integrals of products of functions, instead of probabilities of intersections of events. The results of the present paper extend similarly.

PROOF. Let A_n be a sequence mixing with density p. By diagonal construction one obtains a subsequence C_n of A_n such that

$$\mathbf{P}(C_n/QM) \to p \tag{13}$$

$$(Q=\Omega,\ Q=A_{n_1}\dots A_{n_t};\ n_1,\dots,n_t=1,2,\dots;\ t=1,2,\dots).$$

Let now ε_n be a sequence of positive numbers. We choose positive integers $m_1<m_2<\cdots$ so large that

$$p-\varepsilon_i<\mathbf{P}(C_{m_i}/QM)<p+\varepsilon_i \tag{14}$$

$$(Q=\Omega,\ Q=C_{n_1}\dots C_{n_t};\ n_1,\dots,n_t=m_1,\dots,m_t;\ t=1,\dots,i-1;\ i=1,2,\dots).$$

Let $B_n=C_{m_n}$ $(n=1,2,\dots)$. Since

$$\mathbf{P}(B_{n_1}\dots B_{n_r}/M)=\mathbf{P}(B_{n_1}/M)\,\mathbf{P}(B_{n_2}/B_{n_1}M)\dots\mathbf{P}(B_{n_r}/B_{n_1}\dots B_{n_{r-1}}M), \tag{15}$$

it follows that

$$\prod_{i=1}^{r}(p-\varepsilon_{n_i})<\mathbf{P}(B_{n_1}\dots B_{n_r}/M)<\prod_{i=1}^{r}(p+\varepsilon_{n_i}) \tag{16}$$

$$(n_1<\dots<n_r;\ n_1,\dots,n_r=1,2,\dots;\ r=1,2,\dots)$$

which implies (12) if the ε_n decreases to zero sufficiently fast.

Applying either Theorem 3 and the diagonal construction, or Theorem B [2], one obtains: if a sequence is mixing with density p, then it is equidistributed of all orders with the same density. We assert more.

THEOREM 4. *For any positive integer r, a sequence of events A_n is equidistributed (r,p) if and only if it is mixing with density p.*

We prove here the "only if" part of the theorem. We require from [2] a definition and a lemma (reformulation of Lemma 2 [2]).

DEFINITION 4. A sequence of events A_n is called *condensed* $(\mathbf{P},r,p)$, if for some $\delta>0$

$$\mathbf{P}(A_{n_1}\dots A_{n_r})\geqq p^r+\delta \qquad (n_1,\dots,n_r=1,2,\dots). \tag{17}$$

A sequence containing a $(\mathbf{P},r,p)$ condensed subsequence is called $(\mathbf{P},r,p)$ *semicondensed.*

LEMMA 3. *If a sequence of events A_n is $(\mathbf{P},r,p)$ semicondensed, then it is $(\mathbf{P},s,p)$ semicondensed for all integers $s>r$.*

Let A_n be a sequence of events equidistributed (r,p), we shall show that A_n is mixing with density p. Assume this not to be true, then for some non-null event M and some subsequence B_n of A_n

$$\lim \mathbf{P}(B_n/M)=q\neq p. \tag{18}$$

Denote the conditional probability given M by $\mathbf{P}_M$; if $q>p$, then the

sequence A_n is $(\mathbf{P}_M, 1, p)$ semicondensed, hence, by Lemma 3, it is $(\mathbf{P}_M, r, p)$ semicondensed which contradicts the assumption that the sequence A_n is (r, p) equidistributed. If $q < p$, we consider a subsequence C_n of B_n such that for some $\varepsilon > 0$

$$\mathbf{P}(C_n/M) < p - \varepsilon \qquad (n = 1, 2, \ldots), \tag{19}$$

then a subsequence D_n of C_n and an $\varepsilon' > 0$ such that

$$\mathbf{P}(D_{n_1} \ldots D_{n_r}/M) > p^r - \varepsilon' \qquad (n_1, \ldots, n_r = 1, 2, \ldots). \tag{20}$$

If ε' is chosen sufficiently small, it follows from the relation

$$\mathbf{P}(D_1 D_{n_2} \ldots D_{n_r}/M) = \mathbf{P}(D_1/M)\,\mathbf{P}(D_{n_2} \ldots D_{n_r}/D_1 M) \tag{21}$$
$$(n_2, \ldots, n_r = 1, 2, \ldots)$$

that the sequence D_n is $(\mathbf{P}_{D_1 M}, r-1, p)$ condensed, hence, by Lemma 3, the sequence A_n is $(\mathbf{P}_{D_1 M}, r, p)$ semicondensed which again contradicts the assumption that the sequence A_n is (r, p) equidistributed. Theorem 4 is proven.

4. Mixing of higher orders. Let Ω_0 be the interval $[0, 1]$, $\mathcal{O}_0$ the class of Lebesgue measurable sets of Ω_0, $\mathbf{P}_0$ the Lebesgue measure on $\mathcal{O}_0$. We have the following theorem:

THEOREM 5. *If a sequence of events A_n is mixing (r, p), then it is mixing (s, p) for all positive integers $s < r$. If the "measure space" $[\Omega, \mathcal{O}, \mathbf{P}] = [\Omega_0, \mathcal{O}_0, \mathbf{P}_0]$,*[3] *then for each p $(0 < p < 1)$ and for each positive integer r, there exist sequences mixing (r, p) and not mixing $(r+1, p), (r+2, p), \ldots$.*

PROOF. To prove the first part of the theorem it is enough to show that a sequence mixing (r, p) is mixing $(r-1, p)$. Let a sequence A_n be mixing (r, p); by Theorem 4, A_n is mixing with density p. For arbitrary non-null event M, $\mathbf{P}(A_{n_1} \ldots A_{n_r}/M)$ can be made arbitrarily close to p^r and $\mathbf{P}(A_{n_r}/A_{n_1} \ldots A_{n_{r-1}} M)$ arbitrarily close to p, if at first $n_1 < \cdots < n_{r-1}$ are chosen large, then for $n_1, \ldots, n_{r-1}$ fixed, n_r is chosen large. It follows that $\mathbf{P}(A_{n_1} \ldots A_{n_{r-1}}/M)$ can be made arbitrarily close to p^{r-1} if $n_1 < \cdots < n_{r-1}$ are chosen large; the sequence A_n is $(r-1, p)$ mixing.

To prove the second part of the theorem we need a lemma.

LEMMA 4. *If $[\Omega, \mathcal{O}, \mathbf{P}] = [\Omega_0, \mathcal{O}_0, \mathbf{P}_0]$, then for each p $(0 < p < 1)$ and for each positive integer r, there exist events $A_1, \ldots, A_{r+1}$ such that*

$$\mathbf{P}(A_1 \ldots A_{r+1}) \neq p^{r+1}, \quad \mathbf{P}(A_{n_1} \ldots A_{n_t}) = p^t \tag{22}$$
$$(n_1 < \ldots < n_t;\ n_1, \ldots, n_t = 1, \ldots, r;\ t = 1, \ldots, r).$$

Lemma 4 may be proved by induction on r, the induction step con-

[3] This condition is not very restrictive. Compare Theorem C ([7], p. 173).

sisting in the formation of a product of measure spaces and application of the Transferring Principle ([6], p. 261).

Let now

$$[\Omega_i, \mathcal{A}_i, \mathbf{P}_i] = [\Omega_0, \mathcal{A}_0, \mathbf{P}_0] \qquad (i = 1, 2, \ldots). \tag{23}$$

Consider the infinite-dimensional product space

$$[\Omega_\infty, \mathcal{A}_\infty, \mathbf{P}_\infty] = \left[\mathop{\times}_{i=1}^{\infty} \Omega_i, \mathop{\times}_{i=1}^{\infty} \mathcal{A}_i, \mathop{\times}_{i=1}^{\infty} \mathbf{P}_i \right] \tag{24}$$

isomorphic to $[\Omega_0, \mathcal{A}_0, \mathbf{P}_0]$ ([7], p 157; [6], p. 261). Let $A_1, \ldots, A_{r+1}$ be sets of $\mathcal{A}_0$ satisfying (22). We define $A_{m,n} \in \mathcal{A}_\infty$ by

$$A_{m,n} = \Omega_1 \times \cdots \times \Omega_{m-1} \times A_n \times \Omega_{m+1} \times \cdots \tag{25}$$
$$(m = 1, 2, \ldots;\ n = 1, \ldots, r+1).$$

It is easy to see that the sequence

$$A_{1,1}, \ldots, A_{1,r+1}, A_{2,1}, \ldots, A_{2,r+1}, A_{3,1}, \ldots \tag{26}$$

is mixing (r, p), but not $(r+1, p)$, $(r+2, p)$, ...

We note at the end that between mixing with density p, i. e. equidistribution (r, p), and mixing (r, p), there is perhaps place for yet another notion.

DEFINITION 5. A sequence of events A_n is called semimixing of order r and with density p, or semimixing (r, p), if each subsequence of A_n contains a further subsequence B_n such that for each non-null event M

$$\lim_{\substack{n_1 \to \infty \\ n_1 < \cdots < n_r}} \mathbf{P}(B_{n_1} \ldots B_{n_r}/M) = p^r. \tag{27}$$

The example of the sequence (26) shows that semimixing (s, p) does not imply mixing (s, p) for $s \geqq 2$. We do not know if for $s \geqq 2$ semimixing (s, p) is implied by mixing with density p.

(Received 24 February 1960)

References

[1] A. RÉNYI, On mixing sequences of sets, *Acta Math. Acad. Sci. Hung.*, **9** (1958), pp. 215—228.
[2] L. SUCHESTON, On sequences of events of which the probabilities admit a positive lower limit, *Journal London Math. Soc.*, **34** (1959), pp. 386—394.
[3] G. G. LORENTZ, Remark on a paper by Visser, *Journal London Math. Soc.*, **35** (1960), pp. 205—208.
[4] K. MAURIN, *Metody przestrzeni Hilberta* (Warszawa, 1959).
[5] F. P. RAMSEY, On a problem in formal logic, *Proc. London Math. Soc.* (2), **30** (1930), pp, 264—286.
[6] B. JESSEN, The theory of integration in a space of an infinite number of dimensions, *Acta Math.*, **63** (1934), pp. 249—323.
[7] P. R. HALMOS, *Measure theory* (New York, 1950).

ÜBER EIN REGULÄRES PROBLEM DER VARIATIONSRECHNUNG

Von

A. KÓSA (Budapest)

(*Vorgelegt von* A. RÉNYI)

In [2] haben wir folgendes Problem untersucht:

Es sei n eine beliebige positive ganze Zahl und G ein beliebiges Gebiet in der Ebene. Nehmen wir von der Grundfunktion $f(x, y, y', \ldots, y^{(n)})$ an, daß sie auf der Menge

$$R = \{(x, y) \in G, y^{(i)} \text{ beliebig endlich für } i = 1, \ldots, n\}$$

alle partiellen Ableitungen bis zur 2-ten Ordnung inklusive besitzt. Wir nehmen zwei zu G gehörende Punkte (x_1, y_1), (x_2, y_2), $x_1 \neq x_2$, und geben $2(n-1)$ beliebige Zahlen $y_1^{(i)}, y_2^{(i)}$ $(i = 1, \ldots, n-1)$ an. Wir definieren jetzt die Funktionenklasse $D^{(n)}$ folgendermaßen: $y(x) \in D^{(n)}$ im Intervall $[x_1, x_2]$, wenn $y(x)$ dort stetig ist, und sich das Intervall $[x_1, x_2]$ in eine endliche Anzahl von Teilintervallen zerlegen läßt derart, daß $y(x)$ im Inneren jedes Teilintervalls n-mal stetig differenzierbar ist, und jede Ableitung (bis zur n-ten Ordnung inklusive) in den Teilungspunkten endlichen rechts- und linksseitigen Grenzwert hat. Im folgenden werden wir irgendeinen Punkt $x_0 \in (x_1, x_2)$ Eckenpunkt der Funktion $y(x) \in D^{(n)}$ nennen, wenn die Ungleichung $y^{(i)}(x_0 - 0) \neq y^{(i)}(x_0 + 0)$ wenigstens für ein i $(i = 1, \ldots, n)$ besteht.

Wir definieren nachher die zulässige Funktionenklasse M folgendermaßen: $y(x) \in M$, wenn 1. $y(x) \in D^{(n)}$ im Intervall $[x_1, x_2]$, 2. $y^{(i)}(x_1) = y_1^{(i)}$, $y^{(i)}(x_2) = y_2^{(i)}$ für $i = 0, 1, \ldots, n-1$, 3. $(x, y(x)) \in G$ für $x_1 \leqq x \leqq x_2$. Für alle Funktionen $y(x) \in M$ hat die Operation

$$I[y(x)] = \int_{x_1}^{x_2} f(x, y(x), y'(x), \ldots, y^{(n)}(x))\, dx \tag{1}$$

einen bestimmten endlichen Wert.

In Bezug auf das obige Problem haben wir in [2] gezeigt, daß, wenn die Operation (1) bei der Funktion $y(x) \in M$ ein starkes relatives Minimum[1] annimmt, sie dann im Intervall $[x_1, x_2]$ dem System von Differentialgleichungen

$$f_{y^{(i)}}(x, y, y', \ldots, y^{(n)}) = 0 \quad \text{für} \quad i = 2, \ldots, n \tag{2}$$

[1] Die Definition des starken relativen Minimums s. in [2], S. 27—28.

und — die Eckenpunkte ausgenommen — der Gleichung

$$f_y - \frac{d}{dx} f_{y'} = 0 \tag{3}$$

genügen soll.[2]

Die Gleichungen des Systems (2) sind gewöhnliche Differentialgleichungen n-ter Ordnung, die Gleichung (3) geht aber nur dann in eine gewöhnliche Differentialgleichung $(n+1)$-ter Ordnung über, wenn die das Minimum liefernden Funktionen wenigstens $(n+1)$-mal differenzierbar sind. In Verbindung damit entsteht die Frage: ist es möglich eine Klasse der Probleme so auszuwählen, daß die zu dieser Klasse gehörenden Operationen nur bei wenigstens $(n+1)$-mal stetig differenzierbaren Funktionen ein Minimum annehmen können?

Es ist bekannt, daß im Falle $n=1$ die sogenannten regulären Operationen die entsprechende Eigenschaft besitzen, d. h. eine reguläre Operation kann nur bei 2-mal stetig differenzierbaren Funktionen ein Minimum annehmen.[3] Im Falle $n=1$ gelangt man zum Begriff der regulären Operation folgendermaßen: da die das Minimum liefernde Funktion immer der Legendreschen Bedingung $(f_{y'y'}(x, y(x), y'(x)) \geqq 0$ für $x_1 \leqq x \leqq x_2)$ genügen soll, werden diejenigen Operationen, welche die Legendresche Bedingung *bei allen zulässigen Funktionen* so erfüllen, daß das Gleichheitszeichen überhaupt nicht vorkommt, d. h. bei welchen $f_{y'y'}(x, y, y') > 0$ ist, regulär genannt.

Den Begriff der regulären Operation übertragen wir auf beliebiges positives ganzes n folgendermaßen: in [2] wurde gezeigt, daß, wenn die Operation (1) bei der Funktion $y(x) \in M$ ein starkes relatives Minimum annimmt, die Funktion $y(x)$ für beliebige Zahlen $\sigma_1, \ldots, \sigma_n$ die Ungleichung

$$\sum_{i,j=1}^{n} f_{y^{(i)}y^{(j)}}(x, y(x), y'(x), \ldots, y^{(n)}(x))\sigma_i\sigma_j \geqq 0 \tag{4}$$

in jedem Punkte des Intervalls $[x_1, x_2]$ erfüllen soll. Offensichtlich reduziert sich (4) im Falle $n=1$ eben auf die Legendresche Bedingung. Der obigen Überlegung entsprechend führen wir die folgende Definition ein:

Wir nennen die Operation (1) *regulär, wenn die quadratische Form*

$$\sum_{i,j=1}^{n} f_{y^{(i)}y^{(j)}}(x, y, y', \ldots, y^{(n)})\sigma_i\sigma_j \tag{5}$$

auf der Menge R positiv definit ist.

[2] S. [2], S. 28 und 36.
[3] S. z. B. [1], S. 161.

Nachher kann folgender Satz festgestellt werden:

SATZ I. *Wenn die Operation* (1) *regulär ist und bei der Funktion* $y(x) \in M$ *ein starkes relatives Minimum annimmt, ist* $y(x)$ *wenigstens* $(n+1)$-*mal stetig differenzierbar.*[4]

Die Beweisführung wird in zwei Schritten durchgeführt.

a) Wir zeigen zuerst, daß, wenn eine solche Funktion $y(x) \in M$ existiert, welche in irgendeinem Punkte $x_0 \in (x_1, x_2)$ einen Eckenpunkt besitzt, und bei welcher die Operation (1) ein starkes relatives Minimum annimmt, die Operation (1) nicht regulär sein kann. Mit anderen Worten bedeutet das, daß eine reguläre Operation ein starkes relatives Minimum nur bei wenigstens n-mal stetig differenzierbaren Funktionen annehmen kann.

Nehmen wir also an, daß die Operation (1) bei der Funktion $y(x) \in M$ ein starkes relatives Minimum annimmt, und die Funktion $y(x)$ im Punkte x_0 einen Eckenpunkt besitzt. Dann sollen die sog. Eckenbedingungen

$$\begin{aligned} &f_{y'}(x_0, y(x_0), y'(x_0-0), \ldots, y^{(n)}(x_0-0)) = \\ &= f_{y'}(x_0, y(x_0), y'(x_0+0), \ldots, y^{(n)}(x_0+0)) \end{aligned} \tag{6}$$

und

$$\begin{aligned} &f(x_0, y(x_0), y'(x_0-0), \ldots, y^{(n)}(x_0-0)) - \\ &- y'(x_0-0) \cdot f_{y'}(x_0, y(x_0), y'(x_0-0), \ldots, y^{(n)}(x_0-0)) = \\ &= f(x_0, y(x_0), y'(x_0+0), \ldots, y^{(n)}(x_0+0)) - \\ &- y'(x_0+0) \cdot f_{y'}(x_0, y(x_0), y'(x_0+0), \ldots, y^{(n)}(x_0+0)) \end{aligned} \tag{7}$$

bestehen.[5]

Wir führen die folgenden Bezeichnungen ein:

$$a_i = y^{(i)}(x_0-0,) \quad b_i = y^{(i)}(x_0+0), \quad \sigma_i = b_i - a_i \quad \text{für} \quad i = 1, \ldots, n.$$

Nach Voraussetzung ist

$$\sum_{i=1}^{n} \sigma_i^2 \neq 0. \tag{8}$$

Betrachten wir jetzt die Funktion

$$\begin{aligned} F(t) &= f(x_0, y(x_0), a_1 + t\sigma_1, \ldots, a_n + t\sigma_n) - \\ &- t \sum_{i=1}^{n} \sigma_i f_{y^{(i)}}(x_0, y(x_0), a_1 + t\sigma_1, \ldots, a_n + t\sigma_n). \end{aligned} \tag{9}$$

[4] Der Satz ist ohne Beweis in [3] zu finden.
[5] S. [2], S. 28 und 46.

Unter Berücksichtigung der Bedingungen (2), (6) und (7) erhält man nach einer einfachen Rechnung

$$F(0)=F(1).$$

Nach dem Rolleschen Satz existiert dann ein solches ω $(0<\omega<1)$, für welches

(10) $$F'(\omega)=0.$$

Wir führen die Differentiation durch:

(11)
$$\begin{aligned}
F'(\omega)=\sum_{i=1}^{n}\sigma_i f_{y^{(i)}}(x_0, y(x_0), a_1+\omega\sigma_1, \dots, a_n+\omega\sigma_n)-\\
-\sum_{i=1}^{n}\sigma_i f_{y^{(i)}}(x_0, y(x_0), a_1+\omega\sigma_1, \dots, a_n+\omega\sigma_n)-\\
-\omega\sum_{i,j=1}^{n} f_{y^{(i)}y^{(j)}}(x_0, y(x_0), a_1+\omega\sigma_1, \dots, a_n+\omega\sigma_n)\sigma_i\sigma_j=\\
=-\omega\sum_{i,j=1}^{n} f_{y^{(i)}y^{(j)}}(x_0, y(x_0), a_1+\omega\sigma_1, \dots, a_n+\omega\sigma_n)\sigma_i\sigma_j=0.
\end{aligned}$$

Da $\omega\neq0$, erhalten wir unter Berücksichtigung von (8), daß die behandelte Operation nicht regulär ist.

Damit haben wir den ersten Teil unserer Behauptung bewiesen.

b) Es ist noch übrig nachzuweisen, daß, wenn die Operation (1) regulär ist, und sie bei einer n-mal stetig differenzierbaren Funktion $y(x)\in M$ ein starkes relatives Minimum annimmt, $y^{(n+1)}(x)$ im ganzen Intervall $[x_1, x_2]$ existiert und dort stetig ist. Im Falle $n=1$ ist dieser Teil unserer Behauptung als Hilbertscher Satz bekannt.[6] Wir führen jetzt den Nachweis für $n\geqq2$ durch.

Zu diesem Zwecke betrachten wir die implizite Funktion

(12) $$\varphi(x, u)=f_{y^{(n)}}(x, y(x), y'(x), \dots, y^{(n-1)}(x), u)=0.$$

Nach (2) ist die Funktion

(13) $$u=y^{(n)}(x)$$

eine Lösung der Gleichung (12) im Intervall $[x_1, x_2]$. Da die Operation regulär ist, gilt die Ungleichung

$$\varphi_u(x, y^{(n)}(x))=f_{y^{(n)}y^{(n)}}(x, y(x), y'(x), \dots, y^{(n)}(x))>0$$

im ganzen Intervall $[x_1, x_2$.

[6] S. z. B. [1], S. 144.

Daraus folgt auf Grund des wohlbekannten Satzes über die impliziten Funktionen unter Berücksichtigung der Voraussetzung für die Grundfunktion, daß die Lösungsfunktion (13) im ganzen Intervall $[x_1, x_2]$ stetig differenzierbar ist.

Damit haben wir unsere Behauptung völlig bewiesen.

Nachher ist es leicht, den folgenden Satz zu beweisen:

SATZ II. *Nehmen wir an, daß die Operation* (1) *regulär ist und bei der Funktion* $y(x) \in M$ *ein starkes relatives Minimum annimmt; ferner sei die Grundfunktion* $f(x, y, y', \ldots, y^{(n)})$ *auf der Menge* R $(2+k)$*-mal stetig differenzierbar. In diesem Falle ist* $y(x)$ *im Intervall* $[x_1, x_2]$ *wenigstens* $(n+k+1)$*-mal stetig differenzierbar.*

Wir betrachten zuerst den Fall $n \geqq 2$. Die Funktion $y(x)$ soll nach (2) der Gleichung

$$f_{y^{(n)}}(x, y(x), y'(x), \ldots, y^{(n)}(x)) = 0 \quad \text{für} \quad x_1 \leqq x \leqq x_2 \tag{14}$$

genügen. Früher haben wir gezeigt, daß $y(x)$ wenigstens $(n+1)$-mal stetig differenzierbar ist. Durch Differentiation bekommen wir aus (14) folgendes

$$\begin{gathered} f_{y^{(n)}x}(x, y(x), y'(x), \ldots, y^{(n)}(x)) + \\ + \sum_{i=0}^{n} f_{y^{(n)}y^{(i)}}(x, y(x), y'(x), \ldots, y^{(n)}(x)) y^{(i+1)}(x) = 0. \end{gathered} \tag{15}$$

Da die Operation regulär ist, können wir (15) in folgender Form schreiben:

$$y^{(n+1)}(x) =$$

$$= -\frac{f_{y^{(n)}x}(x, y(x), y'(x), \ldots, y^{(n)}(x)) + \sum_{i=0}^{n-1} f_{y^{(n)}y^{(i)}}(x, y(x), y'(x), \ldots, y^{(n)}(x)) \cdot y^{(i+1)}(x)}{f_{y^{(n)}y^{(n)}}(x, y(x), y'(x), \ldots, y^{(n)}(x))}.$$

Daraus folgt unsere Behauptung unmittelbar.

Für $n = 1$ kann man den Beweis ähnlicherweise durchführen, man soll nur statt der Gleichung (14) die aus (3) entstehende Euler—Lagrangesche Differentialgleichung nehmen.

(Eingegangen am 16. März 1960.)

Literaturverzeichnis

[1] G. A. Bliss, *Calculus of variations* (Chicago, 1925).

[2] A. Kósa, Notwendige Bedingungen für die diskontinuierlichen Lösungen von den Variationsproblemen n-ter Ordnung, *Acta Math. Acad. Sci. Hung.*, **11** (1960), S. 23—48.

[3] A. Kósa, Conditions nécessaires se rapportant aux problèmes discontinus d'ordre supérieur dans le calcul des variations, *Atti della Acc. Naz. dei Lincei*, (1960), S. 328—331.

ÜBER EIN PROBLEM VON G. ALEXITS

Von

K. TANDORI (Szeged)

(*Vorgelegt von* G. ALEXITS)

§ 1. Einleitung

G. ALEXITS hat den folgenden Satz bewiesen:[1]

Es sei $\{q_k\}$ *eine positive, monoton nichtwachsende Zahlenfolge. Bestehen*

$$\sum_{k=1}^{\infty} \frac{q_k}{\sqrt{k}} < \infty \tag{1}$$

und

$$c_k = O(q_k), \tag{2}$$

dann ist die Orthogonalreihe

$$\sum_{k=1}^{\infty} c_k \varphi_k(x) \tag{3}$$

fast überall $(C, 1)$*-summierbar.*

Das folgende Problem kann gestellt werden: Folgt aus (1) und (2), daß die Orthogonalreihe (3) fast überall sehr stark summierbar ist, d. h. eine Funktion $f(x) \in L^2$ gibt derart, daß für jede Indexfolge $\nu_1 < \cdots < \nu_n < \cdots$

$$\sum_{n=1}^{N} [f(x) - s_{\nu_n}(x)]^2 = o(N)$$

fast überall besteht, wo $s_m(x)$ die m-te Partialsumme der Orthogonalreihe (3) bezeichnet?

G ALEXITS vermutete, daß die Antwort auf dieses Problem negativ ist. Wir werden diese Vermutung beweisen. Es gilt nämlich der folgende Satz:

Es gibt ein im Grundintervall $[a, b]$ *orthonormiertes Funktionensystem* $\{\Phi_k(x)\}$, *eine positive, monoton nichtwachsende Koeffizientenfolge* $\{a_k\}$ *mit*

$$\sum_{k=1}^{\infty} \frac{a_k}{\sqrt{k}} < \infty \tag{4}$$

und eine Indexfolge $\nu_1 < \cdots < \nu_n < \cdots$ *derart, daß die Folge*

$$\left\{ \frac{S_{\nu_1}(x) + \cdots + S_{\nu_n}(x)}{n} \right\} \tag{5}$$

[1] G. ALEXITS, Ein Summationssatz für Orthogonalreihen, *Acta Math. Acad. Sci. Hung.*, 7 (1956), S. 5—9.

m $[a, b]$ *fast überall divergiert, wo* $S_m(x)$ *die m-te Partialsumme der Orthogonalreihe*

$$\sum_{k=1}^{\infty} a_k \Phi_k(x)$$

bezeichnet.

Dieses Resultat ist eine Verschärfung eines von mir bewiesenen vorherigen Satzes.[2]

§ 2. Beweis des Satzes

Zum Beweis benötigen wir den folgenden Hilfssatz:

HILFSSATZ. *Es seien* $p(\geqq 4), a_1, \ldots, a_p$ *gegebene natürliche Zahlen,* $t_0 = 0, t_i = a_1 + \cdots + a_i$ *für* $i = 1, \ldots, p$ *und* $\bar{a}_l = a_i$ *für* $t_{i-1} < l \leqq t_i$ $(i = 1, \ldots, p)$. *Dann kann ein orthonormiertes Funktionensystem von Treppenfunktionen*[3] $h_l(x) = h_l(p; \{a\}; x)$ $(l = 1, \ldots, t_p)$ *im Grundintervall* $[a, b]$ *angegeben werden, so daß die folgende Bedingung erfüllt wird:*

Es gibt eine einfache Menge[4] $E = E(\{a\}) (\subseteqq [0, 1])$ *vom Maße* $\operatorname{mes}(E) = \frac{1}{5}$ *mit der Eigenschaft, daß für jedes* $x \in E$ *eine von x abhängige natürliche Zahl* $m(x) (< p)$ *gibt derart, daß die Funktionenwerte* $h_1(x), \ldots, h_{t_{m(x)}}(x)$ *nichtnegativ sind und*

$$\frac{1}{\sqrt{\bar{a}_1}} h_1(x) + \cdots + \frac{1}{\sqrt{\bar{a}_{t_{m(x)}}}} h_{t_{m(x)}}(x) \geqq A \sqrt{p} \log p$$

gilt, wo A eine von x und p unabhängige, positive Konstante bedeutet.

Dieser Hilfssatz ist bekannt.[5]

Durch vollständige Induktion werden wir drei Indexfolgen $\{p_n\}$ $(p_0 = 0,\ 4 \leqq p_1 < \cdots < p_n < \cdots)$; $\{M_n\}$ $(M_n \geqq 1, n = 1, 2, \ldots)$; $(0 \leqq) \nu_1 < \cdots \cdots < \nu_n < \cdots$ definieren, mit folgender Eigenschaften:

$$\sum_{n=1}^{\infty} \frac{1}{\log p_n} < \infty; \tag{6}$$

$$2^{p_1 + \cdots + p_n - 1} = M_n \qquad (n = 1, 2, \ldots); \tag{7}$$

[2] K. TANDORI, Über die orthogonalen Funktionen. IV (Starke Summation), *Acta Sci. Math. Szeged,* **19** (1958), S. 18—25, Satz II.

[3] D. h. für jede Funktion $h_l(x)$ kann das Intervall [0,1] in endlich viele Teilintervalle zerlegt werden derart, daß $h_l(x)$ in jedem Teilintervall konstant ist.

[4] D. h. E ist die Vereinigung endlich vieler Intervalle.

[5] K. TANDORI, Über die orthogonalen Funktionen. II (Summation), *Acta Sci. Math. Szeged,* **18** (1957), S. 149—168, Hilfssatz III.

für $1 \leqq i \leqq p_n$ ist

(8) $$\nu_{2^{p_0+\cdots+p_{n-1}+i}} = M_0 p_0 + \cdots + M_{n-1} p_{n-1} + i M_n \qquad (n = 1, 2, \ldots).$$

Sei in der Tat $p_1 = 4$, $M_1 = 2^3$ und $\nu_1 = 0$, $\nu_{2^i - j} = iM_1 - j$ für $1 \leqq i \leqq p_1$, $j = 0, \ldots, 2^{i-1} - 1$, ferner $N(\geqq 1)$ eine natürliche Zahl. Wir nehmen an, die $p_1, \ldots, p_N, M_1, \ldots, M_N$ und $\nu_1, \ldots, \nu_{2^{p_0+\cdots+p_N}}$ seien schon definiert, so daß (7) und (8) für $n = 1, \ldots, N$ erfüllt sind und für $n = 1, \ldots, N$

(9) $$\log p_n \geqq 2^n$$

besteht. Wir setzen dann $p_{N+1} = 2^{2^{N+1}}$ und $\nu_{2^{p_0+\cdots+p_N+i}-j} = M_0 p_0 + \cdots$ $\cdots + M_N p_N + i M_{N+1} - j$ für $1 \leqq i \leqq p_{N+1}$, $j = 0, \ldots, 2^{i-1} = 1$. (Nach der Definition von M_{N+1} ist dies möglich.) Dann ist auch (8) für $n = N+1$ erfüllt. Auf diese Weise erhalten wir die geforderten Indexfolgen, für welche also (7), (8) und (9) bei jedem n bestehen. Aus (9) folgt dann, daß auch (6) erfüllt ist.

Durch vollständige Induktion werden wir eine positive, monoton nichtwachsende Koeffizientenfolge $\{a_k\}$, ein in $[a, b]$ orthonormiertes System von Treppenfunktionen $\{\Phi_k(x)\}$ und eine Folge von einfachen Mengen $\{F_n\}$ $(F_n \subseteq [a, b])$ mit folgenden Eigenschaften definieren:

Die Mengen F_n sind stochastisch unabhängig[6] und für jedes $n(\geqq 1)$ gilt

(10) $$\operatorname{mes}(F_n) = \frac{b-a}{5};$$

es gilt ferner (4), und für $x \in F_n$ gibt es einen von x abhängigen Index $\mu(x)(< p_n)$, so daß

(11) $$\left| S_{\nu_{2^{p_0+\cdots+p_{n-1}+\mu(x)}}}(x) - S_{\nu_{2^{p_0+\cdots+p_{n-1}}}}(x) \right| \geqq \frac{A}{\sqrt{b-a}}.$$

Es sei nämlich

$$a_k = \frac{1}{\sqrt{M_1 p_1}\,\log p_1} \quad (1 \leqq k \leqq M_1 p_1), \quad a_k = \frac{1}{\sqrt{M_n p_n}\,\log p_n}$$

für $p_1 M_1 + \cdots + p_{n-1} M_{n-1} < k \leqq p_1 M_1 + \cdots + p_n M_n$ $(n = 2, 3, \ldots)$. Offensichtlich ist die Koeffizientenfolge $\{a_k\}$ positiv, monoton nichtwachsend, und auf Grund von (6) ist die Beziehung (4) erfüllt.

Für $p = p_1$, $a_1^{(1)} = \cdots = a_{p_1}^{(1)} = M_1$ wenden wir unseren Hilfssatz an. Es sei

$$\Phi_k(x) = \frac{1}{\sqrt{b-a}} h_k([a, b]; p_1; \{a^{(1)}\}; x) \qquad (k = 1, \ldots, M_1 p_1)$$

[6] Betreffs dieser Definition siehe z. B. K. Tandori, Über die orthogonalen Funktionen, *Acta Sci. Math. Szeged*, 18 (1957), S. 57—130.

und
$$F_1 = E([a, b]; \{a^{(1)}\}).^7$$
Offensichtlich sind diese Funktionen Treppenfunktionen, sie bilden im $[a, b]$ ein orthonormiertes Funktionensystem, F_1 ist eine einfache Menge, für $n=1$ besteht (10) und auf Grund von (8) folgt, daß für $n=1$ auch (11) besteht.

Es sei $N(\geqq 1)$ eine natürliche Zahl. Wir nehmen an, daß die Funktionen $\Phi_k(x)$ $(1 \leqq k \leqq M_0 p_0 + \cdots + M_N p_N)$ und die Mengen $F_1, \ldots, F_N$ schon definiert sind derart, daß diese Funktionen Treppenfunktionen sind, im $[a, b]$ ein orthonormiertes System bilden, die Mengen $F_1, \ldots, F_N$ einfach und stochastisch unabhängig sind, und für $n = 1, \ldots, N$ die Beziehungen (10) und (11) erfüllt sind.

Dann kann das Intervall $[a, b]$ in endlich viele Teilintervalle I_r $(1 \leqq r \leqq \varrho)$ zerlegt werden, so daß jede Funktion $\Phi_k(x)$ $(1 \leqq k \leqq M_0 p_0 + \cdots + M_N p_N)$ in jedem I_r konstant ist und jede Menge F_n $(1 \leqq n \leqq N)$ die Vereinigung einiger I_r ist. Die zwei Hälften von I_r werden mit I_r' bzw. I_r'' bezeichnet. Wir wenden den Hilfssatz mit $p = p_{N+1}$, $a_1^{(N+1)} = \cdots = a_{p_{N+1}}^{(N+1)} = M_{N+1}$ an. Es seien
$$\Phi_{M_0 p_0 + \ldots + M_N p_N + l}(x) = \frac{1}{\sqrt{b-a}} \sum_{r=1}^{\varrho} \big(h_l(I_r'; p_{N+1}; \{a^{(N+1)}\}; x) -$$
$$- h_l(I_r''; p_{N+1}; \{a^{(N+1)}\}; x)\big) \qquad (l = 1, \ldots, M_{N+1} p_{N+1})$$
und
$$F_{N+1} = \bigcup_{r=1}^{\varrho} \big(E(I_r'; \{a^{(N+1)}\}) \cup E(I_r''; \{a^{(N+1)}\})\big).$$

Offensichtlich sind die Funktionen $\Phi_k(x)$ mit $M_0 p_0 + \cdots + M_N p_N < k \leqq M_0 p_0 + \cdots + M_{N+1} p_{N+1}$ Treppenfunktionen, und F_{N+1} ist eine einfache Menge. Durch einfache Rechnung kann man einsehen, daß die Funktionen $\Phi_k(x)$ mit $1 \leqq k \leqq M_0 p_0 + \cdots + M_{N+1} p_{N+1}$ in $[a, b]$ ein orthonormiertes System bilden und die Mengen $F_1, \ldots, F_{N+1}$ stochastisch unabhängig sind. Es folgt leicht, daß für $n = N+1$ auch die Beziehungen (10) und (11) bestehen.

Wir haben somit durch vollständige Induktion ein Funktionensystem $\{\Phi_k(x)\}$ und eine Mengenfolge $\{F_n\}$ mit den erwähnten Eigenschaften erhalten.

Aus der stochastischen Unabhängigkeit der Mengen F_n und aus (10) ergibt sich durch Anwendung des zweiten Borel—Cantellischen Lemmas:

[7] Für ein endliches Intervall $I = [u, v]$ sei
$$f(I; x) = \begin{cases} f\left(\dfrac{x-u}{v-u}\right) & \text{für} \quad u < x < v, \\ 0 & \text{sonst,} \end{cases}$$
und $H(I)$ bezeichne die Bildmenge von H bei der Transformation $y = (v-u)x + u$.

mes $(\overline{\lim\limits_{n\to\infty}} F_n) = b-a$. Ist $x \in \overline{\lim\limits_{n\to\infty}} F_n$, so ist (11) für unendlich viele n erfüllt; d. h. die Folge $\{S_{\nu_{2n}}(x)\}$ divergiert fast überall.

Da aus (4)

$$\sum_{k=1}^{\infty} a_k^2 < \infty$$

folgt, so ergibt sich auf Grund eines meiner vorherigen Resultate,[8] daß die Folge (5) fast überall divergiert.

Damit haben wir unseren Satz vollständig bewiesen.

(Eingegangen am 27. April 1960.)

[8] S. a. a. O. [2], Hilfssatz.

Technikai szerkesztő: Molnár Ferenc

A kiadásért felel az Akadémiai Kiadó igazgatója Műszaki szerkesztő: Farkas Sándor
A kézirat beérkezett: 1960. VI. 15. — Terjedelem: 20 (A/5) ív, 16 ábra

Szegedi Nyomda Vállalat 60-2079

The Acta Mathematica publish papers on mathematics in English, German, French and Russian.

The Acta Mathematica appear in parts of various size, making up volumes.

Manuscripts should be adressed to:

Acta Mathematica, Budapest 502, Postafiók 24.

Correspondence with the editors and publishers should be sent to the same address

The rate of subs.cription to the Acta Mathematica is 110 forints a volume. Orders may be placed with ,,*Kultura*'' Foreign Trade Company for Books and Newspapers (Budapest, I., Fő utca 32. Account No. 43-790-057-181) or with representatives abroad.

Les Acta Mathematica paraissent en français, allemand, anglais et russe et publient des mémoires du domaine des sciences mathématiques.

Les Acta Mathematica sont publiés sous forme de fascicules qui seront réunis en volumes

On est prié d'envoyer les manuscrits destinés à la rédaction à l'adresse suivante:

Acta Mathematica, Budapest 502, Postafiók 24.

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 110 forints par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur de Livres et Journaux ,,*Kultura*'' (Budapest, I, Fő utca 32. Compte-courant No. 43-790-057-181) ou à l'étrange chez tous les représentants ou dépositaires.

,,*Acta Mathematica*" публикует трактаты из области математических наук на русском немецком, английском и французском языках.

,,*Acta Mathematica*" выходит отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рукописи следует направлять по адресу:

Acta Mathematica, Budapest 502, Postafiók 24.

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена ,,*Acta Mathematica*" — 110 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет ,,*Kultura*" (*Budapest, I., Fő utca 32.* Текущий счет № 43-790-057-181) или его заграничные представительства и уполномоченные.

Reprinted by arrangement with the publishers
„KULTURA" Hungarian Trading Company
for Books and Newspapers
Budapest, POB. 149.
Hungary

ACTA MATHEMATICA
ACADEMIAE SCIENTIARUM HUNGARICAE
Том XI — Вып. 1—2

РЕЗЮМЕ

О ЛИНЕЙНЫХ ОТОБРАЖЕНИЯХ ИНТЕГРАЛЬНОГО ТИПА ПРОСТРАНСТВА НЕПРЕРЫВНЫХ ВЕКТОРНЫХ ФУНКЦИЙ

И. Зингер (Бухарест)

Пусть Q — компактное пространство, E и F — банаховые пространства, $C(Q, F)$ — пространство всех непрерывных отображений Q в F, и E' — сопряженное к E пространство. Доказывается следующая

Теорема 1. Каждое интегральное [2] линейное отображение u пространства $C(Q, F)$ в пространство E' является мажорированным [1] и

$$\|\mu\| \leqq \|u\|_{\mathrm{int}},$$

где через μ обозначается наименьшая мажоранта u, а через $\|u\|_{\mathrm{int}}$ интегральная норма u. Для того, чтобы обратное было верным, т. е. для того, чтобы каждое мажорированное линейное отображение $C(Q, F)$ в E' было интегральным, необходимо и достаточно, чтобы каноническое линейное отображение J пространства $F \hat{\otimes} E$ в $F \check{\otimes} E$ было изоморфизмом $F \hat{\otimes} E$ на $F \check{\otimes} E$. В этом случае имеем

$$\|u\|_{\mathrm{int}} \leqq \|J^{-1}\| \, \|\mu\|.$$

В частности, если $F = R$, получается вновь теорема 1 статьи [6].

О КОНЕЧНЫХ ФОРМУЛАХ СУММИРОВАНИЯ АРИФМЕТИЧЕСКОГО ХАРАКТЕРА. II

Л. Карлиц (Дурхам, США)

Примыкая к работам [2] и [3], автор в находящейся в печати статье [1] доказал, что для любой системы периодических с периодом 1 функций $f_1(x), \ldots, f_n(x)$, удовлетворяющих функциональному уравнению типа (1.1), тождественно выполняется (1.2), где $a_1, \ldots, a_n$ попарно относительно простые натуральные числа, $A = a_1 a_2 \ldots a_n$ и $C_k^{(a_k)}$ зависящая лишь от k и a_k постоянная. В случае $k \to \infty$ (1.2) переходит в соответствующее интегральное соотношение; самый важный частный случай: $n = 2$ и $f_1(x) = \bar{B}_p(x), f_2(x) = \bar{B}_q(x)$. (Здесь $\bar{B}_n(x)$ суть функции Бернулли, определенные формулами

$$\frac{te^{tx}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x)\, t^n \quad (|t| < 2\pi); \quad \bar{B}_n(x) = B_n(x) \quad (0 \leqq x < 1), \quad \bar{B}_n(x+1) = \bar{B}(x).$$

Настоящая работа изучает случай многочленов Эйлера и обобщенных функций Эйлера $\Phi_m(x, \zeta)$ $\left(\zeta = e^{\frac{2\mu\pi i}{\nu}},\ (\mu, \nu) = 1\right)$ (см. (1.3)—(1.5)).

Доказываются следующие теоремы:

Теорема I. Пусть ζ_j есть примитивный ν_j-ый корень из единицы, а $f_j(x, \zeta_j)$ $(j=1, 2, \ldots, n)$ система функций, для элементов которых выполняются соотношения (1.8), (1.9) при постоянных коэффициентах $C_j^{(k)}$ и $k \equiv 1 \pmod{\nu}$, $\nu = [\nu_1, \ldots, \nu_n]$ ($[\nu_1, \ldots, \nu_n]$ обозначает наименьшее общее кратное.) Пусть далее $a_1, \ldots, a_n$ такие попарно относительно простые положительные целые, для которых $a_j \equiv 1 \pmod{\nu}$ $(j=1, 2, \ldots, n)$, и $A = a_1 a_2 \ldots a_n$. Тогда выполняется (1.12), где $C = C_1^{(a_1)} \ldots C_n^{(a_n)}$.

Теорема II. Пусть ζ_j те же, что и выше, η_j такой $\lambda_j \nu_j$-ый примитивный корень из единицы, для которого $\zeta_j = \eta_j^{\lambda_j}$. Предположим, что функции $f_j(x, \zeta_j)$ $(j=1, 2, \ldots, n)$ удовлетворяют (1.8) и (1.9) и

$$\sum_{r=0}^{\lambda_j - 1} \eta_j^r f_j\left(x + \frac{r}{\lambda_j}, \zeta_j\right) = K^{(\lambda_j)} g_j(\lambda_j x, \eta_j) \qquad (j = 1, 2, \ldots, n)$$

с постоянными коэффициентами $K^{(\lambda_j)}$. Пусть далее $a_1, a_2, \ldots, a_n$ попарно относительно простые числа, удовлетворяющие условиям $a_j | \lambda_j$, $a_j \equiv 1 \pmod{\lambda \nu}$ $(j=1, 2, \ldots, n)$, где $\lambda = [\lambda_1, \ldots, \lambda_n]$, $\nu = [\nu_1, \ldots, \nu_n]$. Тогда при обозначениях $A = a_1 a_2 \ldots a_n$, $K = K^{(a_1)} \ldots K^{(a_n)}$ и $k \equiv 1 \pmod{\lambda \nu}$ имеет место (5.4).

НЕОБХОДИМЫЕ УСЛОВИЯ, ОТНОСЯЩИЕСЯ К НЕГЛАДКИМ РЕШЕНИЯМ ВАРИАЦИОННОЙ ПРОБЛЕМЫ n-ОГО ПОРЯДКА

А. Коша (Будапешт)

В настоящей работе изучаются относящиеся к операции

$$I[y] = \int_{x_1}^{x_2} f(x, y, y', \ldots, y^{(n)})\, dx$$

необходимые условия в том случае, когда о допустимых функциях предполагается, что они сами непрерывны, но их производные — вплоть до n-ой — лишь кусочно-непрерывны.

Пусть $y(x)$ $(x_1 \leqq x \leqq x_2)$ любая допустимая функция. Точки, где одна из производных $y'(x), \ldots, y^{(n)}(x)$ терпит разрыв, назовем угловыми точками и введем функцию

$$E(x, y, y', \ldots, y^{(n)}; p', \ldots, p^{(n)}) =$$
$$= f(x, y, p', \ldots, p^{(n)}) - f(x, y, y', \ldots, y^{(n)}) - (p' - y') f_{y'}(x, y, y', \ldots, y^{(n)}).$$

Для относительного слабого минимума получены следующие необходимые условия:

I. $f_{y^{(i)}} = 0$, если $x_1 < x < x_2$ $(i = 2, \ldots, n)$;

II. $f_y - \dfrac{d}{dx} f_{y'} = 0$, если $x_1 < x < x_2$, за исключением угловых точек;

III. $f_{y^{(i)} y^{(i)}} \geqq 0$, если $x_1 \leqq x \leqq x_2$ $(i = 1, \ldots, n)$;

IV. $f_{y'}|_{x-0} = f_{y'}|_{x+0}$, если $x_1 < x < x_2$;

V. существует такое $\varepsilon > 0$, что

$$E(x, y(x), y'(x), \ldots, y^{(n)}(x); p', \ldots, p^{(n)}) \geqq 0$$

в (x_1, x_2) за исключением угловых точек, если

$$|p^{(i)} - y^{(i)}(x)| < \varepsilon \qquad (i = 1, \ldots, n).$$

В случае относительного сильного минимума к предыдущим прибавляются еще следующие условия:

VI. $E(x, y(x), y'(x), \ldots, y^{(n)}(x);\ p', \ldots, p^{(n)}) \geqq 0,$ если $x_1 \leqq x \leqq x_2;\ p', \ldots, p^{(n)}$ любые;

VII. $f - y' f_{y'}|_{x-0} = f - y' f_{y'}|_{x+0},$ если $x_1 < x < x_2$.

При выводе условий V и VI было использовано уравнение II, а при выводе последнего система уравнений I.

Если допустимая функция $y(x)$ удовлетворяет условию V (или более сильному условию VI), то она удовлетворяет и условиям I и III ($i \geqq 2$). Мы получили, таким образом, следующий существенный результат: в случае $n = 1$ хорошо известное уравнение Эйлера—Лагранжа (I), условие Лежандра (III, $i = 1$), необходимое условие Вейерштрасса (VI) и условия Вейерштрасса—Эрдманна (II, VII) являются необходимыми условиями не только для случая $n = 1$, но и для разрывных проблем вышесказанного типа.

ЗАМЕЧАНИЯ ОБ ИНТЕРПОЛИРОВАНИИ

О. Киш (Будапешт)

Автор дает явный вид многочленов $R_n(z)$ не выше $rn - 1$-ой степени, удовлетворяющих в узлах

$$z_k = \exp i \frac{2\pi}{n} k \qquad (k = 1, 2, \ldots, n)$$

условиям

$$R_n^{(m)}(z_k) = \alpha_{k,\,m} \qquad (n \geqq r,\ r \geqq 2,\ 1 \leqq k \leqq n,\ m = 0, 1, \ldots, r-2, r).$$

В случае $r = 2$ и $r = 3$ исследуется сходимость этих многочленов к определенной при $|z| \leqq 1$ функции $f(z)$, если

$$\alpha_{k,\,0} = f(z_k) \qquad (k = 1, 2, \ldots, n).$$

ОБ АФФИННОЙ СВЯЗИ СТЕПЕННЫХ ПРОСТРАНСТВ КАСАЮЩИХСЯ ПРОСТРАНСТВ

Л. Тамашши (Дебрецен)

В некоторой точке пространства тензоры одинакового порядка образуют пространство произведений касающегося пространства. Аффинное сопоставление касающихся пространств соседних точек, обладающее обычным свойством (обычная аффинная связь векторов), уже индуцирует аффинную связь пространств произведений. Но непосредственное сопоставление пространств произведений дает более общую аффинную связь между этими пространствами. Работа занимается изучением этой связи.

Такая связь определяется некоторым дифференциальным оператором. В § 1 определяется наиболее общий дифференциальный оператор, удовлетворяющий условиям (1. 1)—(1. 6), и, таким образом, аффинная связь. В § 2 изучается, когда эта связь сводится к обычной. В § 3 дается условие эквивалентности таких пространств. В § 4 находятся величины кривизны и кручения и изучается их геометрическое значение. Нэконец, в § 5 рассматривается геометрическая сущность случая редукции.

ОДИН ПРИНЦИП ДВОЙСТВЕННОСТИ В МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

К. Шаркади (Будапешт)

Работа обобщает принцип двойственности, рассмотренный в работах Штейнхауса и Одерфелда ([18], [9], [10]). Этот принцип в сущности эквивалентен уравнению (1); здесь распределение $\varkappa_n$ биномиальное с параметрами n, ζ и ζ также случайная величина, априорное распределение которой равномерно на отрезке (0, 1). Обобщение, данное для гипергеометрического, полигипергеометрического, полиномиального распределения и распределения Пуассона, доказывается с помощью моделей. В разделе 6 автор предлагает улучшение метода качественного контроля Одерфелда, основанного на принципе двойственности: метода априорного распределения типа β [9].

ЗАМЕЧАНИЕ К ОДНОЙ ТЕОРЕМЕ ЧИНКВИНИ

А. Шарма (Лукноу, Индия)

С. Чинквини [1], [2], [3] и П. Монтел [4], [5] распространили теорему о среднем, относящуюся к вещественным функциям, на функции комплексной переменной. Автор, используя метод Чинквини, доказывает следующую теорему:

Если функция $f(z)$ регулярна в круге $|z-\alpha| \leqq r$ и $f''(\alpha) \neq 0$, то существует окружность $|z-\alpha| = r_0' < r$, для которой выполняется равенство

$$\begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ f(z_0) & f(z_1) & f'\left(\frac{z_0+z_1}{2}\right) \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ z_0 & z_1 & 1 \\ z_0^3 & z_1^3 & 3\left(\frac{z_0+z_1}{2}\right)^2 \end{vmatrix} \frac{f'''(\xi)}{3!},$$

если $|z_0-\alpha| = |z_1-\alpha| = r_0 < r_0'$ и $\alpha = \dfrac{z_0+z_1}{3}$. Здесь ξ некоторая точка, лежащая внутри окружности $|z-\alpha| = r_0$.

О ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЕ ДЛЯ СУММ СЛУЧАЙНОГО ЧИСЛА СЛАГАЕМЫХ

А. Реньи (Будапешт)

В работе доказывается следующая

Теорема. Пусть ξ_n $(n=1, 2, \ldots)$ независимые и одинаково распределенные случайные величины, с математическим ожиданием 0 и дисперсией 1. Пусть

$$\eta_n = \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{\sqrt{n}}.$$

Пусть ν_n $(n=1, 2, \ldots)$ последовательность случайных величин, принимающих лишь целые положительные значения, и предположим, что $\frac{\nu_n}{n}$ стремится при $n \to \infty$ по вероятности к λ, где λ — положительная случайная величина с дискретным распределением.

Тогда имеет место

$$\lim_{n\to\infty} \mathbf{P}(\eta_{\nu_n} < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du.$$

Доказательство опирается кроме метода Дэблина на одну теорему автора (см. [5]).

НЕРАВЕНСТВА, ОТНОСЯЩИЕСЯ К ПЕРИМЕТРУ, ПЛОЩАДИ И МОМЕНТУ ИНЕРЦИИ ВЫПУКЛЫХ КРИВЫХ

Х. Сакс (Халле)

Пусть $\mathfrak{C}$ есть замкнутая кривая в n-мерном евклидовом пространстве R_n, L — ее длина; $I = \int\limits_{\mathfrak{C}} r^2\, ds$ — момент инерции кривой относительно ее тяжести; Q — квадратичное среднее расстояний между точками P, P', т. е.

$$Q = \left\{ \frac{1}{L^2} \int\limits_{\mathfrak{C}}\int\limits_{\mathfrak{C}} |PP'|^2\, ds\, ds' \right\}^{\frac{1}{2}}.$$

Далее, если $\mathfrak{C}$ плоская кривая, то обозначим ее площадь через F, если $\mathfrak{C}$ выпуклая, то пусть $J = \int\limits_{\mathfrak{C}} r^2\, d\vartheta$ — кривизной момент инерции относительно кривизной точки тяжести Штейнера. Пусть, далее, $\mathfrak{K}$ обозначает окружность, $\mathfrak{D}$ — равносторонний треугольник, а $\mathfrak{N}$ — кривую, состоящую из двух точек некоторого отрезка и его концов („игла", „Nadel"). $A \underset{\mathfrak{K}}{\geqq} B$ означает следующее: $A \geqq B$ и равенство имеет место только в случае окружности.

Между L, I и Q выполняется следующее (встречающееся уже у Штейнера) соотношение:

$$LQ^2 = 2I. \tag{10}$$

Имеют место следующие неравенства:

a) для любой $\mathfrak{C}$

(2) $L^3 - 4\pi^2 I \underset{\mathfrak{K}}{\geqq} 0$ ([6], стр. 125, теорема 3);

b) для плоских кривых $\mathfrak{C}$

(4) $L^3 - 8\pi LF + 4\pi^2 I \underset{\mathfrak{K}}{\geqq} 0$ (настоящая работа, ср. (3)),

(10a) $LI - 4F^2 \underset{\mathfrak{K}}{\geqq} 0$ ([5], стр. 49);

c) для выпуклых кривых $\mathfrak{C}$

(12) $F \underset{\mathfrak{N}}{\geqq} 0$ (очевидно),

(13) $54I - L^3 \underset{\mathfrak{D}}{\geqq} 0$ ([7], стр. 127, теорема 1),

(14) $8\pi^2 LI - L^4 - (4\pi F)^2 \underset{\mathfrak{K}}{\geqq} 0$ (настоящая работа, § 2),

(18) $2\pi F + \pi J - L^2 \underset{\mathfrak{K}}{\geqq} 0$ (настоящая работа, § 2)

(20) $\pi L^2 - 8J \underset{\mathfrak{N}}{\geqq} 0$ ([8], стр. 347, теорема 5).

Из этих неравенств легко получить дальнейшие изопериметрические неравенства, а именно

a) для любых кривых $\mathfrak{C}$: (2′), (24);

b) для плоских кривых $\mathfrak{C}$: (5), (6), (6′), (9), (9′), (15′);

c) для выпуклых кривых $\mathfrak{C}$: (13′), (14′), (14a—e), (16), (16′), (17), (18′), (19), (21), (21′), (22), (23), (25), (26). Для выпуклых кривых $\mathfrak{C}$ из (2) и (14) получаются еще (4) и (10a).

В настоящей работе как новые результаты доказываются (14) и (18), используя неравенство Шварца как основное вспомогательное средство. Для звездообразных плоских кривых (14) не всегда имеет место.

Можно задать такие числа $L > 0$, F и I, которые удовлетворяют неравенствам (2), (12), (13), (14) (а, следовательно, и (4) и (10a)) со знаком > 0, но, несмотря на это, нет такой выпуклой кривой, длина которой L, площадь F и наименьший момент инерции I.

Из (4) получается следующая, заслуживающая внимания

Теорема. Для каждой выпуклой конфигурации, ограниченной кривой $\mathfrak{C}$ длины 2π и содержащей центр O внутри себя, выполняется соотношение

(1) $$\int_{\mathfrak{C}} r^2\, ds \geqq \int_{\mathfrak{C}} r^2\, d\varphi$$

(r, ψ: полярные координаты). Равенство имеет место лишь в том случае, если $\mathfrak{C}$ есть окружность радиуса 1 с центром в центре системы координат. (См. фиг. 1.)

ЗАМЕЧАНИЯ К ТЕОРИИ АБЕЛЕВЫХ ГРУПП. II

Л. Фукс (Будапешт)

В первой части работы автор обобщает введенную Куликовым базисную подгруппу на любые абелевы группы. Для некоторого простого числа p под p-базисной подгруппой любой абелевой группы G понимается такая подгруппа B, которая 1. является прямой суммой циклических групп порядка степени p и бесконечных, 2. p-сервантна в том смысле, что $p^r B = B \cap p^r G$ $(r = 1, 2, \dots)$, 3. факторгруппа G/B p-полна, т. е. $p(G/B) = G/B$.

Теорема. Группа G содержит для всех p p-базисную подгруппу и относящиеся к одному и тому же p p-базисные подгруппы все изоморфны.

Теорема имеет ряд приложений.

Во второй части работы изучаются группы, все (абелевы) пополнения которых группами без кручения расщепляемы. Эти группы являются обобщениями алгебраически компактных групп и их важность очевидна и из того факта, что группа пополнений, $\mathrm{Ext}(L, K)$ всегда обладает этим свойством.

ОБОБЩЕННАЯ АССОЦИАТИВНОСТЬ И БИСИММЕТРИЯ В КВАЗИГРУППАХ

Я. Ацел (Дебрецен), В. Д. Белоусов (Бельцы, СССР) и М. Хоссу (Мишколц)

Авторы дают общее решение функциональных уравнений (1) и (2) в любых квазигруппах. Они доказывают, что если некоторое множество под четырьмя или шестью действиями образует квазигруппу и для любых трех или четырех элементов выполняется (1) или (2) соответственно, то все четыре квазигруппы изотопны некоторой группе, а все шесть — группе Абеля, причем общее решение имеет вид (3) и (16) соответственно. Авторы рассматривают также специальный случай вещественных непрерывных функций, вопросы единственности и приложения, в первую очередь к теории сетей. Они дают несколько отличающееся от обычного обобщение последних для квазигрупп с помощью теорем Томсена.

ПРОБЛЕМЫ СТРУКТУРЫ ПУТИ ПРИ БЛУЖДАНИИ

П. Эрдёш (Будапешт) и С. И. Тейлор (Бирмингем)

В настоящей статье авторы занимаются проблемами блужданий на прямой, в плоскости и вообще в d-мерном пространсве. В случае $d = 2$ доказывается, что

$$\lim_{n=\infty} \mathbf{P}(R_n < x \log n) = 1 - e^{-\pi x},$$

где R_n обозначает число возвращений случайно блуждающей на плоскости точки в

исходную точку после n шагов. При $d=1$ и $d=2$ они определяют аналог теоремы повторного логарифма для очень частых и очень редких возвращений и получают теоремы о распределении моментов возвращения. В случае $d \geqq 3$ они изучают теоремы, связанные со скоростью стремления в бесконечность, здесь не удается полностью перенести наиболее острую теорему, соответствующую закону повторного логарифма. Наконец, в случае $d \geqq 2$ они исследуют, каково число точек, через которые случайно блуждающая точка пройдет k раз за n шагов, а также исследуют максимальные пути, но здесь получается сильная теорема лишь при $d \geqq 3$.

МЕТОД ДЛЯ ТОЧНОЙ ОЦЕНКИ ПОГРЕШНОСТИ ПРИ ПРИБЛИЖЕННОМ ВЫЧИСЛЕНИИ МНОГОКРАТНЫХ ИНТЕГРАЛОВ

Л. Ч. Хсу (Чангчунь, Китай)

Автор изучает процесс $L(f)$, определенный формулой (2) для элементов f класса функций $W = W^{(m,n)}(M, M_\mu, N_\nu; Q)$, определенного формулой (1), который относится к приближенному определению интеграла

$$\iint_D f(x, y)\,dx\,dy,$$

распространенного на область $D \in Q$. Он находит точную верхнюю грань

$$\varepsilon(L, D) = \sup_{f \in W} |\delta(f, L, D)|$$

погрешности

$$\delta(f, L, D) = \iint_D f(x, y)\,dx\,dy - L(f).$$

Она дается формулой (20), где $F(t_1, t_2)$, $G_\nu(t)$ и $H_\mu(t)$ определяются формулой (19) и для натуральных r

$$K_r(u) = \begin{cases} \dfrac{1}{(r-1)!}\,u^{r-1}, & \text{если} \quad u \geqq 0, \\ 0, & \text{если} \quad u < 0. \end{cases}$$

К ТЕОРИИ КОМИТАНТОВ

С. Голомб (Краков) и М. Кухаржевски (Катовице)

Основной результат работы — доказательство того, что дважды ковариантный тензор не имеет такой функции (комитанта), значение которой p раз контравариантный и q раз ковариантный тензор, если $p+q$ нечетное число. Результат легко может быть обобщен на тот случай, когда исходный тензор r раз контравариантен и s раз ковариантен и $r+s$ четное число.

О КОВАРИАНТНОМ ДИФФЕРЕНЦИАЛЕ, ОБРАЗОВАННОМ ИЗ g_{ik}

А. Моор (Сегед)

Согласно теореме 1 работы из симметричного тензора g_{ik} в случае выполнения подходящих условий непрерывности могут быть образованы лишь классические параметры сдвига $\Gamma_i{}^j{}_k$. Затем работа изучает некоторые типы ковариантных дифференциалов, которые могут быть образованы из g_{ik}. Получены следующие основные результаты: Согласно теореме 2 общий ковариантный дифференциал контравариантного вектора ξ^i зависит лишь от величин ξ^i, g_{ij} и

$$\nabla_k \xi^i \overset{\text{def}}{=} \partial_k \xi^i + \Gamma_r{}^i{}_k \xi^r.$$

$\nabla_k \xi^i$ есть как раз классический ковариантный дифференциал. Общий ковариантный дифференциал ковариантного вектора η_i согласно теореме 4 является функцией от η_i, g_{ij} и $\nabla_k \eta_i$, где $\nabla_k \eta_i$ снова означает классический ковариантный дифференциал. Дальнейшие теоремы изучают, может ли общий ковариантный дифференциал не зависеть от величин g_{ab} и $\partial_c g_{ab}$. В ковариантном дифференциале контравариантного вектора эти величины всегда должны фигурировать, но в случае ковариантных векторов нет, так как уже само $\partial_{[i} \eta_{k]}$ удовлетворяет условиям, определяющим ковариантный дифференциал.

О ГЕОМЕТРИЧЕСКОМ ПРИЛОЖЕНИИ ОДНОЙ СИСТЕМЫ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Л. Тамашши (Дебрецен)

В настоящей работе изучается геометрическое приложение систем дифференциальных уравнений типа (1. 1). Система дифференциальных уравнений такого типа встречается при изучении эквивалентности дифференциально-геометрических пространств. Уравнения образуют формулы преобразования геометрических объектов, определяющие структуру пространства (или их соответствующие преобразованные). В § 1 доказывается, что, хотя в случае метрических пространств разрешимости этой системы дифференциальных уравнений достаточно для эквивалентности, в случае неметрических пространств может случиться, что система дифференциальных уравнений разрешима, а два пространства все же не эквивалентны. Причина этого заключается в том, что система функций, являющаяся решением дифференциального уравнения, не обязательно обратима и тогда эти функции не могут определить преобразование (сохраняющее размерность). В § 2 дается достаточное условие эквивалентности аффинно-связных пространств, связанное с числом параметров, встречающихся в решении системы дифференциальных уравнений. Число параметров может быть определено и без решения системы дифференциальных уравнений. Найденное здесь число с некоторой точки зрения является наилучшим. В § 3 дается система дифференциальных уравнений типа (1. 1), разрешимости которой необходимо и достаточно для эквивалентности двух аффинно-связных пространств. С помощью условий их интегрируемости вопрос об эквивалентности может быть разрешен разрешимостью некоторой обыкновенной системы уравнений.

РАЗЛОЖИМОСТЬ ДИФФЕРЕНЦИАЛЬНО-ГЕОМЕТРИЧЕСКИХ ПРОСТРАНСТВ

О. Варга (Будапешт)

В работе, прежде всего, определяется представление финслерова пространства F_n в виде произведения пространств F_r и F_{n-r}, обобщая соответствующее понятие римановой геометрии. Основным результатом работы является теорема 1, в которой дается необходимое и достаточное условие представимости финслерова пространства в виде произведения пространств.

ACTA MATHEMATICA
ACADEMIAE SCIENTIARUM HUNGARICAE
Том XI — Вып. 3—4

РЕЗЮМЕ

СКАЛЯРНЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ ИНВАРИАНТЫ ТРЕТЬЕГО ПОРЯДКА РИМАНОВОГО ПРОСТРАНСТВА ТРЕХ ИЗМЕРЕНИЙ

П. И. Петров (Казань, СССР)

Под абсолютным скалярным дифференциальным инвариантом третьего порядка риманового пространства n измерений разумеется такая скалярная функция I, которая зависит от компонентов метрического тензора и их частных производных (до третьего порядка включительно), неизменяющаяся при любых обратимых преобразованиях координат.

Цель статьи — построить базис полной системы абсолютных скалярных дифференциальных инвариантов третьего порядка риманового пространства трех измерений. Основной результат — теорема 1. Опираясь на лемму 2, из нее можно получить теорему 2, относящуюся к базису функционально независимых скалярных дифференциальных инвариантов третьего порядка конформно-плоского риманового пространства трех измерений.

В теореме 3 перечислены метрические инварианты тернарной кубической формы. Теорема 4 является распространением известного результата Суворова и Христоффеля на случай пространств с неопределенным мер-определением.

ОБ ОДНОЙ НОВОЙ ТЕОРЕМЕ ТЕОРИИ ДИОФАНТИЧЕСКОЙ АППРОКСИМАЦИИ

Н. Г. де Брёйн (Амстердам)

П. Туран в книге[1] привел ряд приложений следующей теоремы: Если $b_1, \ldots, b_k$ комплексны,

$$\min_{j=1,\ldots,k} |z_j| = 1$$

и m положительное целое число, то

$$\max_{\nu=m+1,\ldots,m+k} |b_1 z_1^\nu + \cdots + b_k z_k^\nu| \geqq \left(\frac{k}{2e(m+k)}\right)^k |b_1 + \cdots + b_k| \overset{\mathrm{def}}{=} A_{m,k} |b_1 + \cdots + b_k|.$$

В настоящей работе доказывается, что это значение $A_{m,k}$ может быть заменено

[1] P. Turán, *Eine neue Methode in der Analysis und deren Anwendungen* (Budapest, 1953).

большим значением

$$\frac{1}{\sum_{j=0}^{k-1} 2^j \binom{m+j}{j}}$$

и оно является наилучшим для всех m и k. Эту теорему — менее просто — несколько раньше независимо доказал Э. Макаи.

ПОЛУГРУППЫ И КОЛЬЦА С ЛЕВОЙ ЕДИНИЦЕЙ И БЕЗ ЛЕВЫХ ЕДИНИЧНЫХ ЭЛЕМЕНТОВ

Л. Редеи (Сегед)

Элемент ε структуры S называется левой единицей, если в S определено произведение и $\varepsilon S = S$. П. Туран поставил следующий вопрос: Может ли содержать левые единицы кольцо без левых единичных элементов. Хотя и кажется вероятным, что часто встречаются такие кольца, было не легко найти для них пример, фигурирующий в следующей теореме:

Теорема. Полугруппа, определенная уравнениями $\alpha = \alpha^2\beta$, $\beta = \alpha\beta^2$ не содержит левого единичного элемента, но элемент α очевидно является левой единицей. Аналогично обстоит дело с кольцом этой полугруппы, образованным над любой областью целостности.

КЛАССИФИКАЦИЯ ПРЕДСТАВЛЕНИЙ ГРУППЫ

Ф. Лонстра (Гаага, Голландия)

Если α гомоморфизм группы G на группу B, то мы говорим, что $G(\alpha)$ есть представление B. Представления $G(\alpha)$ и $G'(\alpha')$ родственны, если существуют такие гомоморфизмы η и η', что 1. $G\eta \subseteq G'$ и $G'\eta' \subseteq G$, 2. $\alpha = \eta\alpha'$ и $\alpha' = \eta'\alpha$.

Представление $G(\alpha)$ точно тогда родственно идентичному представлению $B(\varepsilon)$, если G расщепляемое расширение ядра гомоморфизма α. Представления свободными группами родственны.

Произведение представлений и частичное упорядочение представлений может быть определено так, что компатибильно с родственностью.

Родственные классы представлений относительно частичного упорядочения образуют полную структуру, где пересечение и соединение может быть характеризованно и с помощью теории групп.

ПРОБЛЕМЫ ПЕРЕСЕЧЕНИЯ ОТНОСИТЕЛЬНО БЛУЖДАНИЙ

П. Эрдёш (Будапешт) и С. И. Тейлор (Бирмингем)

Пусть $\Pi^{(d)}(a, b)$ обозначает путь случайно блуждающей в d-мерном пространстве точки между a-ым и b-ым шагом. Пусть $E_n^{(d)}$ означает событие, что $\Pi^{(d)}(0, n)$ и $\Pi^{(d)}(n+f(n), \infty)$ пересекаются по крайней мере в одной точке. В случае $d = 3$ пусть

$f(n) = n[\varphi(n)]^2$, где $\varphi(n)$ монотонно растет. $E_n^{(3)}$ в том и только в том случае имеет место с вероятностью 1 для бесконечно многих значений n, если

$$\sum_{k=1}^{\infty} \frac{1}{\varphi(2^k)} = \infty.$$

В случае $d = 4$ пусть $f(n) = n\psi(n)$, где $\psi(n)$ монотонно растет. $E_n^{(4)}$ в том и только в том случае имеет место с вероятностью 1 для бесконечно многих значений n, если

$$\sum_{k=1}^{\infty} \frac{1}{k\psi(2^k)} = \infty.$$

При $d > 4$ $E_n^{(d)}$ в том и только в том случае имеет место с вероятностью 1 для бесконечно многих значений n, если

$$\sum_{n=1}^{\infty} \frac{1}{f(n)^{\frac{d-2}{2}}} = \infty.$$

ОБ ОДНОМ СВОЙСТВЕ ЦЕЛЫХ ФУНКЦИЙ

М. Кучма (Краков)

Основной результат работы следующая

Теорема. Если для последовательности α_n с положительными членами

$$\lim_{n\to\infty} \frac{\alpha_n}{n} = A \qquad (0 < A < \infty)$$

и $f(z)$ целая вещественная функция на вещественной оси, то

$$A \varlimsup_{n\to\infty} \sqrt{|f^{(n)}(0)|} \geqq \ln \varlimsup_{n\to\infty} \sqrt[n]{|f(\alpha_n)|}.$$

О ТРИГОНОМЕТРИЧЕСКОМ (0,2)-ИНТЕРПОЛИРОВАНИИ

О. Киш (Будапешт)

Автор доказывает, что при четном n, вообще говоря, не существует тригонометрический многочлен $R_n(x)$ не выше n-ого порядка, удовлетворяющий условиям

$$R_n\left(\frac{2\pi}{n}k\right) = \alpha_k, \quad R_n''\left(\frac{2\pi}{n}k\right) = \beta_k \qquad (k = 0, 1, \ldots, n-1),$$

где α_k и β_k любые вещественные числа. Если n нечетно, то существует единственный тригонометрический многочлен не выше n-ого порядка, удовлетворяющий этим условиям и не содержащий $\sin nx$.

Находится явный вид этих многочленов и исследуется их сходимость к функции $f(x)$, если

$$\alpha_k = f\left(\frac{2\pi}{n}k\right) \qquad (k = 0, 1, \ldots, n-1).$$

О НЕКОТОРЫХ РЕЗУЛЬТАТАХ И ПРОБЛЕМАХ ТЕОРИИ МНОЖЕСТВ

А. Хайнал (Будапешт)

Работа занимается тремя различными группами вопросов теории множеств, изучаемых в разделах 2, 3 и 4, соответственно. Основной результат второго раздела — решение одной проблемы П. Эрдёша и Г. Фодора, поставленной в работе [2]. Он дается теоремой 1, в которой с использованием обобщенной гипотезы континуума доказывается следующее:

Если $\aleph_\alpha$ регулярная мощность и S любое множество мощности $\aleph_{\alpha+1}$, то на S может быть определено такое отображение множеств $f(x)$, которое обладает следующими свойствами:

$\overline{\overline{f(x)}} = \aleph_\alpha$ для всех $x \in S$, $\overline{\overline{f(x) \cap f(y)}} < \aleph_\alpha$ для всех $x, y \in S$ $(x \neq y)$ и не существует свободного множества мощности $\aleph_{\alpha+1}$ (т. е.: если $S' \subset S$ и для любых $x, y \in S'$ $(x \neq y)$, $y \notin f(x)$, $x \notin f(y)$, то $\overline{\overline{S'}} < \aleph_{\alpha+1}$).

В разделе 3 изучается символ „партицио", введенный в [1] П. Эрдёшом и Р. Радо.

Здесь доказываются следующие теоремы: $\omega_1 \nrightarrow (\omega+2, \omega_1)^2$ (теорема 5). $\lambda \to (\omega \cdot n, \alpha)^2$, где α любое счетное порядковое число и n целое число (теорема 6). $\lambda \to (\eta, \alpha \vee \alpha^*)^2$, где α любое счетное порядковое число (теорема 7). $\omega_1 \to (\omega \cdot 2, \omega \cdot n)^2$, где n любое целое число (теорема 8). λ и η типы порядков множеств расположенных по величине вещественных и рациональных чисел, соответственно. При доказательстве теоремы 5 используется гипотеза континуума.

В разделе 4, с использованием обобщенной гипотезы континуума, доказывается следующая теорема (теорема 9): Если $\aleph_\alpha$ любая регулярная мощность и S любое множество мощности $\aleph_{\alpha+1}$, то существует такая система $\mathfrak{S}$ подмножеств множества S, которая обладает следующими свойствами: Если $X \in \mathfrak{S}$, то $\overline{\overline{X}} = \aleph_\alpha$. $\overline{\overline{X \cap Y}} < \aleph_\alpha$, если $X, Y \in \mathfrak{S}$ $(X \neq Y)$. Если $S' \subseteq S$, $\overline{\overline{S'}} = \aleph_{\alpha+1}$, то существует такой $X \in \mathfrak{S}$, для которого $X \subset S'$.

ОБ ОБОСТРЕНИИ НЕКОТОРЫХ „ОДНОСТОРОННИХ" ТЕОРЕМ ТЕОРИИ ДИОФАНТОВЫХ ПРИБЛИЖЕНИЙ

П. Туран (Будапешт)

Автор отмечает, что в качестве новых приложений второй основной теоремы его книги[1] получаются неравенства

$$\int_{Te^{-\frac{\log T}{\sqrt{\log\log T}}}}^{T} R^2(x)\,dx > T^2 e^{-c_2 \frac{\log T \log\log\log T}{\log\log T}},$$

если $T > c_1$, и

$$\int_{Te^{-\frac{\log T}{\sqrt{\log\log T}}}}^{T} [\psi_1(x,k) - \psi_i(x,k)]^2\,dx > T^2 e^{-c_3 \frac{\log T \log\log\log T}{\log\log T}},$$

если $T > \max(c_3, e^{3k^4})$ и $(l, k) = 1$, где c суть явно данные положительные постоянные,

$$R(x) = \sum_{n \leqq x} \Lambda(n)^2 \quad \text{и} \quad \psi_l(x, k) = \sum_{\substack{n \leqq x \\ n \equiv l \bmod k}} \Lambda(n),$$

после чего приступает к „одностороннему" обобщению основных теорем, к которым его снова привели приложения к теории чисел. В этом направлении приводятся следующие две теоремы:

Теорема I. Если для некоторого $0 < k < \frac{\pi}{2}$ $k \leqq |\arc z_j| \leqq \pi$ $(j = 1, \ldots, n)$, m и n натуральные числа, то при $|z_1| \geqq |z_2| \geqq \cdots \geqq |z_n| = 1$

$$\max_{\substack{m+1 \leqq \nu \leqq m+n\left(3+\frac{\pi}{k}\right) \\ \nu \text{ целое}}} \operatorname{Re} \sum_{j=1}^{n} b_j z_j^{\nu} \geqq \left(\operatorname{Re} \sum_{j=1}^{n} b_j\right) \frac{1}{5n} \left(\frac{n}{27(m+n)}\right)^{2n}$$

и

$$\min_{\substack{m+1 \leqq \nu \leqq m+n\left(3+\frac{\pi}{k}\right) \\ \nu \text{ целое}}} \operatorname{Re} \sum_{j=1}^{n} b_j z_j^{\nu} \leqq -\left(\operatorname{Re} \sum_{j=1}^{n} b_j\right) \frac{1}{5n} \left(\frac{n}{27(m+n)}\right)^{2n}.$$

Теорема II. Если с предыдущими k, m и n теперь $1 = |z_1| \geqq |z_2| \geqq \cdots \geqq |z_n|$, то

$$\max_{\substack{m+1 \leqq \nu \leqq m+n\left(3+\frac{\pi}{k}\right) \\ \nu \text{ целое}}} \operatorname{Re} \sum_{j=1}^{n} z_j^{\nu} \geqq \left(\frac{1}{81(m+n)}\right)^{2\left(3+\frac{\pi}{k}\right)n^3}$$

и

$$\min_{\substack{m+1 \leqq \nu \leqq m+n\left(3+\frac{\pi}{k}\right) \\ \nu \text{ целое}}} \operatorname{Re} \sum_{j=1}^{n} z_j^{\nu} \leqq -\left(\frac{1}{81(m+n)}\right)^{2\left(3+\frac{\pi}{k}\right)n^3}.$$

[1] P. Turán, *Eine neue Methode in der Analysis und deren Anwendungen* (Budapest, 1953). Совершенно переработанное английское издание выйдет в серии *Interscience Tracts*.

[2] $\Lambda(n)$ — символ Манголдта: $\log p$, если $n = p^{\alpha}$, p — простое число, 0 для прочих n.

ПРИМЕНЕНИЕ ОДНОГО НОВОГО МЕТОДА СУММИРОВАНИЯ К ТРИГОНОМЕТРИЧЕСКИМ РЯДАМ И РЯДАМ ДИРИХЛЕ

М. Миколаш (Будапешт)

Работа занимается одним новым методом суммирования, так называемым $(\mathfrak{W}_{\pm})$-суммированием, в связи с одной новой теорией производных и интегралов комплексного порядка (см. [9]). Под $(\mathfrak{W}_{\pm})$-суммой ряда

$$A_0 + 2\sum_{n=1}^{\infty}(A_n \cos 2n\pi u + B_n \sin 2n\pi u)$$

(с любыми вещественными или комплексными коэффициентами) в точке $u=x$ понимается предел $A_0 + \lim\limits_{\theta\to+0} \sum_{\pm}(x,\theta)$, где

$$\sum\nolimits_{\pm}(x,\theta) = \sum_{n=1}^{\infty} 2(2n\pi)^{-\theta}\left[A_n \cos\left(2n\pi x \pm \frac{\pi\theta}{2}\right) + B_n \sin\left(2n\pi x \pm \frac{\pi\theta}{2}\right)\right],$$

и считается, что последний ряд сходится при достаточно малых $\theta > 0$.

Метод в случае ряда Фурье, относящегося к отрезку (0,1), измеримой ограниченной функции $f(u)$ делает возможным дать очень простое необходимое и достаточное условие суммируемости, а именно существование

$$f\langle x \pm 0\rangle = \lim_{\theta\to+0}\left(\theta\int_0^{\delta} f(x\pm\theta)\, t^{\theta-1}\, dt\right)$$

в исследуемой точке x. (Здесь $\delta > 0$ можно считать как угодно малым.) $(\mathfrak{W}_{+})$-сумма $f\langle x+0\rangle$ переходит в $f(x+0)$, а $(\mathfrak{W}_{-})$-сумма $f\langle x-0\rangle$ в $f(x-0)$, если функция имеет предел при $u\to x+0$ или $u\to x-0$, соответственно; в любом замкнутом интервале непрерывности имеет место равномерная $(\mathfrak{W}_{\pm})$-суммируемость (теорема I).— Доказывается, что метод $(\mathfrak{W}_{\pm})$ в некотором смысле сильнее исследованных до сих пор методов, так как в правой или левой точке Лебега (см. (4.2)—(4.3)) $f\langle x+0\rangle$ и $f\langle x-0\rangle$, соответственно, всегда существуют, в то время, как $f\langle x+0\rangle$ и $f\langle x-0\rangle$ могут одновременно существовать, без того, чтобы существовало (4.1) (теорема II).[1] Надо отметить, что, в отличие от суммируемости (C) и (A), суммируемость $(\mathfrak{W}_{+})$ или $(\mathfrak{W}_{-})$ зависит лишь от значений функции справа или слева от точки x, соответственно, и что теорема I дает альтернативное решение одной проблемы Фейера (см. [3]): определить в отдельности $f(x+0)$ и $f(x-0)$, и таким образом $f(x+0)-f(x-0)$, из ряда Фурье функции $f(u)$.

После случая дифференцированных рядов Фурье, исследуется $(\mathfrak{W}_{\pm})$-суммирование рядов вида (5.4); здесь предполагается, что $f(u)$ или одна из ее производных принадлежит классу $L^q(0,1)$ $(1 < q < \infty)$. С помощью полученного результата (см. теорему III) и ряды Дирихле могут быть суммируемы в некоторых точках границы полуплоскости сходимости.

[1] Известно, что существование (4.1) наиболее общее достаточное условие суммируемости, фигурирующее в литературе. (См. [6], [14].)

ОХАРАКТЕРИЗОВАНИЕ ПЕРИОДИЧЕСКИХ АБЕЛЕВЫХ ГРУПП ПРИ ВЫБРАННЫХ БАЗИСНЫХ ПОДГРУППАХ

Д. К. Харрисон (Хаверфорд, США)

Пусть S есть прямая сумма циклических p-групп. Под парой групп (f, G) понимается редуцированная p-группа G вместе с таким изоморфизмом f, который отображает S на некоторую базисную подгруппу G. (f_1, G_1) и (f_2, G_2) эквивалентны, если существует такой изоморфизм G_1 на G_2, что $f_2 = gf_1$.

Основной результат работы нахождение соответствия между классами эквивалентности пар групп и замкнутыми подгруппами группы I_S. Здесь $I_S = \mathrm{Hom}\,(Z(p^\infty), S^*|S)$, где S^* обозначает так называемую замкнутую p-группу с базисной подгруппой S.

ОБ ОДНОМ МЕТОДЕ ЧИСЛЕННОГО РЕШЕНИЯ УРАВНЕНИЯ ПУАССОНА В КОНЕЧНЫХ РАЗНОСТЯХ В ЛЮБОЙ ОБЛАСТИ

Е. Эгервари (Будапешт)

Как известно, конечными аналогами линейных дифференциальных уравнений в частных производных являются уравнения в конечных разностях, т. е. специальные системы линейных алгебраических уравнений. Надо заметить, что, несмотря на то, что теорию систем линейных алгебраических уравнений можно считать давно завершенной, известные до сих пор методы решения еще далеко не удовлетворяют во всем требованиям практики. Достаточно, например, отметить, что во многих задачах физики и техники с увеличением числа неизвестных число необходимых действий так быстро растет, что иногда оказывается иллюзорным даже использование мощных быстродействующих математических машин.

В литературе фигурирует ряд методов решения уравнений Пуассона в конечных разностях (и подобных), однако большинство из них может применяться лишь в случае прямоугольной области. В настоящей работе автор выработал метод численного решения уравнения Пуассона в конечных разностях для любой области. Особенности этого метода могут быть охарактеризованы следующим образом.

Матрица коэффициентов, соответствующая конечному оператору Лапласа в случае прямоугольной области может быть партиционирована ко клеточной матрице с переставимыми блокками. Для обращения клеточных матриц такого типа автор выработал алгорифм, с помощью которого порядок матрицы, которую следует обратить, и, таким образом, и число необходимых действий может быть значительно снижено [2]. Любая область, состоящая из точек решетки, всегда может быть включена в прямоугольную область, а из матрицы, обратной к относящейся к этой области матрицы, — которую в силу сказанного уже можно считать известной, — матрица, обратная матрице коэффициентов, относящейся к любой области, может быть вычислена с помощью простых действий, „понижающих ранг" [3]. Предложенный здесь метод ввиду простоты и однотипности действий допускает удобное программирование для автоматических вычислительных машин.

НАИБОЛЕЕ КОРОТКИЕ СЕТИ С ЭЛЕМЕНТАМИ ОДИНАКОВОЙ ПЛОЩАДИ

Л. Фейеш Тот (Будапешт)

В качестве следствия более общей теоремы доказывается: Пусть E есть объединенное множество n граней правильной мозаики с трехреберными вершинами. Среди сетей с прямыми ребрами, разбывающих E на n элементов с равными площадями, правильная является наиболее короткой.

О ВПИСЫВАЕМОМ В ПЛОСКОСТЬ ГРАФЕ ПОЛИГОНА ГАМИЛЬТОНА

В. Т. Татт (Торонто)

Полигоном Гамильтона называется состоящий из граней полигон (простая замкнутая кривая), содержащий все вершины графа. Графциклично k-кратно связный, если опуская менее k граней, его нельзя разбить на две части, каждая из которых содержит полигон. Граф кубичный, если к каждой его вершине примыкают точно три грани. В [2] автор дал вписываемый в плоскость кубичный граф, который не имеет полигона Гамильтона и который циклично 3-кратно связный. Этим он опроверг старую гипотезу Тейта. Настоящая работа содержит конструкцию такого вписываемого в плоскость кубичного графа, который не имеет полигона Гамильтона и который циклично 4-кратно связный.

ОХАРАКТЕРИЗОВАНИЕ НЕКОТОРЫХ КЛАССОВ ФУНКЦИЙ С ПОМОЩЬЮ ТЕОРИИ ПРИБЛИЖЕНИЯ

Д. Кралик (Будапешт)

Пусть $\lambda(x)$ есть такая определенная для положительных значений x положительная, монотонно убывающая к 0 функция, что для некоторого числа $0<\gamma<1$ $\lambda(x)x^{\gamma}$ монотонно неубывает. Мы говорим, что $f(x)\in L^p_\lambda$ $(1\leqq p\leqq+\infty)$, если $f(x)\in L^p\,[0,2\pi]$, $f(x+2\pi)=f(x)$ и

$$\omega_p(\delta,f)=O\left[\lambda\left(\frac{1}{\delta}\right)\right],$$

где

$$\omega_p(\delta,f)=\sup_{|h|\leqq\delta}\left(\int_0^{2\pi}|f(x+h)-f(x)|^p\,dx\right)^{1/p}=\sup_{|h|\leqq\delta}\|f(x+h)-f(x)\|_p.$$

Доказываются следующие теоремы:

Теорема 1. $f(x)\in L^p_\lambda$ в том и только том случае, если

$$\|\sigma_n(x;f)-f(x)\|_p=O[\lambda(n)],$$

где $\sigma_n(x;f)$ n-ая сумма Фейера ряда Фурье функции $f(x)$.

Теорема 2. Пусть $\tilde{f}(x)$ есть функция, сопряженная с $f(x)$. Тогда $f(x)$ и $\tilde{f}(x)$ одновременно $\in L^p_\lambda$. В случае $p=1$ и $p=+\infty$ это можно утверждать лишь если для $\lambda(x)$ выполняется еще соотношение

$$\int_{1/\delta}^{+\infty} \frac{\lambda(x)}{x}\, dx = O\left[\lambda\left(\frac{1}{\delta}\right)\right].$$

Дальнейшие теоремы исследуют возможности локализации теоремы 1 на основании обобщения Н. Бари неравенства Бернштейна для тригонометрических многочленов.

ПРИБЛИЖЕНИЯ С ПОМОЩЬЮ АРИФМЕТИЧЕСКИХ СРЕДНИХ ОБЩИХ ОРТОГОНАЛЬНЫХ РЯДОВ

Г. Алексич и Д. Кралик (Будапешт)

Пусть $\sigma_n(x)$ есть n-ое $(C, 1)$-среднее ряда $\sum\limits_{n=0}^{\infty} c_n \varphi_n(x)$, где $\{\varphi_n(x)\}$ ортонормированная на отрезке $[a, b]$ система функций и $\sum\limits_{n=0}^{\infty} c_n^2 < \infty$. Доказывается, что, если неравенство $\sum\limits_{n=1}^{\infty} c_n^2 \vartheta^2(n) < \infty$ обеспечивает сходимость $\{\sigma_n(x)\}$ на множестве $E \subset [a, b]$ и $\sum\limits_{n=1}^{\infty} c_n^2 \lambda^2(n)\, \vartheta^2(n) < \infty$, где $\lambda(x)$ и $\vartheta(x)$ при $x > 0$ положительные, монотонно стремящиеся к $+\infty$ функции и, кроме того, $\lambda(x)$ направлена вогнутостью вниз и для некоторого $0 < \gamma < 1$ $\frac{x^\gamma}{\lambda(x)}$ монотонно не убывает для достаточно больших x, то для всех $x \in E$

$$\sigma_n(x) - f(x) = o_x\left(\frac{1}{\lambda(n)}\right),$$

где $f(x)$ функция, относящаяся к ряду $\sum\limits_{n=0}^{\infty} c_n \varphi_n(x)$ в силу теоремы Рисса—Фишера. Этим обобщается один более ранний аппроксимационный результат Тандори. Аналогичный результат доказывается относительно ортонормированной на отрезке $[-1, +1]$ системы многочленов, порожденной весом $0 \leq \varrho(x) \leq \frac{\text{Konst.}}{\sqrt{1-x^2}}$, когда играет роль и структура разлагаемой в ряд непрерывной функции. Наконец, обобщается интеграл Вейля дробного порядка и относящаяся к нему теорема Харди и Литлвуда.

Вся статья основывается на следующей общей лемме теории рядов: если $u_1, u_2, \ldots$ суть элементы банахова пространства R и $(C, 1)$-средние σ_n (в метрике R) ряда $\sum\limits_{n=1}^{\infty} u_n$ сходятся к элементу $\sigma \in R$ так, что $\|\sigma - \sigma_n\| = O[\Lambda(n)]$, то $(C, 1)$-средние $\sigma_n(\theta)$ ряда $\sum\limits_{n=1}^{\infty} \theta(n) u_n$ сходятся к элементу $\sigma_n(\theta) \in R$ так, что $\|\sigma_n(\theta) - \sigma(\theta)\| = O[\Lambda(n)\theta(n)]$. Здесь $\Lambda(x)$ и $\theta(x)$ определенные при $x > 0$ положительные, невозрастающие функции, направленные вогнутостью вверх, для которых при достаточно больших x функции $\Lambda(x)x^\gamma$ и $\theta(x)x^\eta$ монотонно неубывают, где γ и η некоторые фиксированные числа, для которых $0 < \gamma < 1$, $0 < \eta < 1$ и $\gamma + \eta < 1$.

О β-РАЗЛОЖЕНИИ ВЕЩЕСТВЕННЫХ ЧИСЕЛ

В. Парри (Лондон)

А. Реньи в работе [1] рассмотрел β-разложение неотрицательного числа x в виде

(1) $$x = \sum_{k=0}^{+\infty} \frac{\varepsilon_k(x)}{\beta^k},$$

где $\beta > 1$, и коэффициенты $\varepsilon_k(x)$ определены следующим образом: $\varepsilon_0(x) = [x]$, $\varepsilon_1(x) = [\beta(x)]$, $\varepsilon_2(x) = [\beta(\beta(x))]$ и т. д., где $[x]$ означает целую часть и (x) дробную часть x, и доказал, что преобразование $T(x) = (\beta x)$ интервала $[0, 1]$ является эргодичным и для всякого $g(x) \in L[0, 1)$ имеет место

(2) $$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(T^k(x)) = M(g),$$

где

(3) $$M(g) = \int_0^1 g(x)\, d\nu_\beta = \frac{\int_0^1 g(x) h_\beta(x)\, dx}{\int_0^1 h_\beta(x)\, dx},$$

т. е. $M(g)$ является средним функции $g(x)$ по мере ν_β, которая является инвариантной по T и абсолютно непрерывной относительно меры Лебега.

Работа занимается эффективным построением меры ν, т. е. функции $h_\nu(x)$. Доказаны следующие теоремы:

Теорема 1. $h_\nu(x)$ удовлетворяет функциональному уравнению

$$\beta h_\beta(x) = \sum_{m=0}^{[\beta - x]} h_\beta\left(\frac{x+m}{\beta}\right).$$

Теорема 2. $h_\beta(x) = \sum_{x < T^n(1)} \frac{1}{\beta_n}$.

Далее изучаются условия, при которых данная последовательность b_n может быть последовательностью коэффициентов β-разложения некоторого числа x (теорема 3).

Число β, коэффициенты β-разложения которого образуют периодическую последовательность, называется β-числом. Число β, которое имеет конечное β-разложение, называется простым β-числом. Доказывается, что β-числа являются алгебраическими, их сопряженые все лежат в круге $|z| < 2$, далее простые β-числа всюду плотны в интервале $1 < \beta < +\infty$ (теорема 5). Далее изучается функция $F(\beta) = \int_0^1 h_\beta(x)\, dx$ (теоремы 6, 7 и 8).

ЗАМЕЧАНИЯ О ПЕРЕМЕШИВАЮЩИХ ПОСЛЕДОВАТЕЛЬНОСТЯХ

Л. Сачестон (Рочестер, США)

Пусть $[\Omega, \mathcal{A}, \mathbf{P}]$ поле вероятностей. Пусть $\mathbf{P}(A|B)$ означает условную вероятность события A при условии B.

Определение 1. А. Реньи [1] назвал последовательность событий $A_n \in \mathcal{A}$ перемешивающей с плотностью p, если для всякого $M \in \mathcal{A}$ с $\mathbf{P}(M) > 0$ имеем

$$\lim_{n\to+\infty} \mathbf{P}(A_n|M) = p, \tag{1}$$

где $0 < p < 1$ и p не зависит от выбора M, и доказал, что если $A_0 = \Omega$ и

$$\lim_{n\to+\infty} \mathbf{P}(A_n|A_k) = p \quad \text{при} \quad k = 0, 1, 2, \ldots, \tag{2}$$

то (1) имеет место.

Автор настоящей статьи ввел (см. [2]) следующие определения:

Определение 2. Последовательность событий $A_n \in \mathcal{A}$ называется (r, p)-равномерно распределенной, если для всякого $M \in \mathcal{A}$, для которого $\mathbf{P}(M) > 0$, во всякой подпоследовательности A_n найдется такая подпоследовательность B_n, что имеет место

$$\lim_{\substack{n_1\to+\infty \\ n_1<n_2<\cdots<n_r}} \mathbf{P}(B_{n_1} \cdots B_{n_r}|M) = p^r. \tag{3}$$

Определение 3. Последовательность событий $A_n \in \mathcal{A}$ называется (r, p) перемешивающей, если

$$\lim_{\substack{n_1\to+\infty \\ n_1<n_2<\cdots<n_r}} \mathbf{P}(A_{n_1} \cdots A_{n_r}|M) = p^r. \tag{4}$$

Автор доказывает, что теорема 2 работы [1] А. Реньи следует из леммы 4 работы [2] автора. Далее доказываются следующие теоремы:

Теорема 4. Для всякого $r = 2, 3, \ldots$ для того, чтобы последовательность A_n была (r, p)-равномерно распределенной, необходимо и достаточно, чтобы A_n была перемешивающей с плотностью p в смысле А. Реньи.

Теорема 5. Если последовательность $\{A_n\}$ является (r, p)-равномерно распределенной, то она является также (s, p)-равномерно распределенной с каждым $s > r$; это не всегда верно для $s < r$.

ОБ ОДНОЙ РЕГУЛЯРНОЙ ПРОБЛЕМЕ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ

А. Коша (Будапешт)

Определим операцию

$$I[y] = \int_{x_1}^{x_2} f(x, y, y', \ldots, y^{(n)})\, dx$$

в таком классе непрерывных функций $y(x)$, где производные $y^{(i)}(x)$ $(i = 1, \ldots, n)$, вообще говоря, лишь кусочно непрерывны. О функции f предполагается, что она дважды непрерывно дифференцируема. Тогда — как было показано в [2] — в случае относитель-

ного сильного минимума одно из необходимых условий лишь тогда переходит в обыкновенное дифференциальное уравнение, если минимизирующие функции по крайней мере $n+1$ раз непрерывно дифференцируемы. Назовем регулярной операцию, для которой квадратичная форма

$$\sum_{i,j=1}^{n} f_{y^{(i)}y^{(j)}}(x, y, y', \ldots, y^{(n)})\, \sigma_i\, \sigma_j$$

положительно определенна для всех допустимых значений. В работе доказывается, что регулярная операция может принимать относительный сильный минимум лишь на $n+1$ раз непрерывно дифференцируемых функциях. Если о функции f предположить, что она обладает производными и выше второго порядка, то в случае регулярной проблемы для минимизирующей функции получаются дальнейшие условия дифференцируемости.

ОБ ОДНОЙ ПРОБЛЕМЕ Г. АЛЕКСИЧА

К. Тандори (Сегед)

Пусть $\{\varphi_n(x)\}$ есть система функций, ортонормированная на отрезке $[a, b]$. Г. Алексич доказал следующую теорему:[1]

Если для положительной, монотонно убывающей последовательности коэффициентов $\{c_k\}$ выполняется условие

$$\sum_{k=1}^{\infty} \frac{c_k}{\sqrt{k}} < \infty, \tag{1}$$

то ортогональный ряд

$$\sum_{k=0}^{\infty} c_k \varphi_k(x) \tag{2}$$

$(C, 1)$-суммируем почти всюду на $[a, b]$.

Ряд (2) называется очень сильно суммируемым на $[a, b]$, если существует такая интегрируемая с квадратом функция $f(x)$, что для любой последовательности индексов $\nu_1 < \cdots < \nu_n < \cdots$ почти всюду на $[a, b]$

$$\sum_{n=1}^{N} [f(x) - s_{\nu_n}(x)]^2 = o(N) \qquad \left(s_n(x) = \sum_{k=0}^{n} c_k \varphi_k(x)\right).$$

Автор, доказывая одну гипотезу Г. Алексича, показывает, что в случае положительной, монотонно убывающей последовательности коэффициентов $\{c_k\}$ из условия (1), вообще говоря, не следует очень сильная суммируемость ряда (2).

Этот результат является обобщением одной предыдущей теоремы автора.[2]

[1] G. Alexits, Ein Summationssatz für Orthogonalreihen, *Acta Math. Acad. Sci. Hung.*, **7** (1956), стр. 5—9.

[2] K. Tandori, Über die orthogonalen Funktionen. IV (Starke Summation), *Acta Sci. Math. Szeged*, **19** (1958), стр. 18—25, Satz 2.